

# Asymptotic Inference for a Linear Stochastic Differential Equation with Time Delay\*

Alexander A. Gushchin  
Steklov Mathematical Institute  
Gubkina 8  
117966 Moscow GSP-1, Russia

Uwe Küchler  
Institut für Mathematik  
Humboldt-Universität zu Berlin  
Unter den Linden 6  
D-10099 Berlin, Germany

May 21, 1997

## Abstract

For the stochastic differential equation

$$dX(t) = \{aX(t) + bX(t-1)\} dt + dW(t), \quad t \geq 0,$$

the local asymptotic properties of the likelihood function are studied. They very depend on the true value of the parameter  $\vartheta = (a, b)^* {}^1$ . Eleven different cases are possible if  $\vartheta$  runs through  $R_2$ . Let  $\hat{\vartheta}_T$  be the maximum likelihood estimator of  $\vartheta$  based on  $(X(t), t \leq T)$ . Applications to the asymptotic behaviour of  $\hat{\vartheta}_T$  as  $T \rightarrow \infty$  are given.

**Keywords:** LAMN; LAN; LAQ; likelihood function; limit theorems for martingales; local asymptotic properties, maximum likelihood estimator; stochastic differential equations; time delay

## 1 Introduction

Assume  $(W(t), t \geq 0)$  is a real-valued standard Wiener process,  $a$  and  $b$  are real numbers and  $(X(t), t \geq -1)$  is a solution of

$$dX(t) = aX(t) dt + bX(t-1) dt + dW(t), \quad t \geq 0, \quad (1.1)$$

---

\* This is a revised version of the Discussion Paper 29/96 (SFB 373) "Asymptotic properties of maximum likelihood estimators for a class of linear stochastic differential equations with time delay". The paper was printed using funds made available by the Deutsche Forschungsgemeinschaft

<sup>1)</sup> \* denotes the transpose of the corresponding vector or matrix

with some fixed initial condition  $X(t) = X_0(t)$ ,  $t \in [-1, 0]$ , where  $X_0(\cdot)$  is a continuous stochastic process independent of  $W(\cdot)$ . The solution  $(X(t), t \geq -1)$  of (1.1) exists, is pathwise uniquely determined and can be represented as

$$X(t) = x_0(t)X_0(0) + b \int_{-1}^0 x_0(t-s-1)X_0(s)ds + \int_0^t x_0(t-s)dW(s), \quad t \geq 0. \quad (1.2)$$

Obviously, it has continuous paths for  $t \geq 0$  with probability one. Here  $(x_0(t), t \geq -1)$  denotes the so-called fundamental solution of the deterministic equation

$$\begin{aligned} \dot{x}(t) &= ax(t) + bx(t-1), & t > 0, \\ x(t) &= 1, & t = 0, \\ x(t) &= 0, & t \in [-1, 0). \end{aligned} \quad (1.3)$$

Equation (1.1) is a very special case of linear stochastic differential equations of the type

$$dX(t) = \int_{-1}^0 X(t+s)a(ds)dt + dM(t), \quad t \geq 0, \quad (1.4)$$

where  $a(\cdot)$  is an arbitrary function of finite variation on  $[-1, 0]$  and  $(M(t), t \geq 0)$  is, e.g., a semimartingale, see Mohammed and Scheutzow (1990).

Assume, the solution  $(X(t), t \in [-1, T])$  of (1.1) for some finite  $T > 0$  has been observed, the parameters  $(a, b)$  are unknown and have to be estimated. Then we have a parametric problem, which generalizes the statistical problem of estimating the parameter in Langevin's equation

$$dX(t) = aX(t)dt + dW(t), \quad t \geq 0 \quad (1.5)$$

(see, e.g., Basawa and Prakasa Rao 1980). Estimation problems for stochastic differential equations with time delay have been considered in few papers up to now, see Dietz (1992) and Küchler and Kutoyants (1996) and the references therein. The model we consider seems to be of interest by the following reasons. First, it is a relatively simple example exhibiting a variety of qualitatively different local asymptotic properties for different values of the parameter. Second, the model already shows some typical effects appearing in estimation problems for equations with time-delayed terms. Third, in contrast to more general delay models, we are able to compute explicitly the rates of convergence and the limit distributions of estimators for every value of the parameter. The solutions of (1.1) form an exponential family of continuous stochastic processes in the sense of Küchler and Sørensen (1989). Thus the maximum likelihood estimator  $\hat{\vartheta}_T$  of  $\vartheta = (a, b)^*$  can be expressed explicitly by

$$\hat{\vartheta}_T = (I_T^0)^{-1} V_T^0,$$

where  $V_T^0$  denotes the vector

$$V_T^0 = \left( \int_0^T X(t) dX(t), \int_0^T X(t-1) dX(t) \right)^*$$

and  $I_T^0$  is the observed Fisher information matrix given by

$$I_T^0 = \begin{pmatrix} \int_0^T X^2(t) dt & \int_0^T X(t)X(t-1) dt \\ \int_0^T X(t)X(t-1) dt & \int_0^T X^2(t-1) dt \end{pmatrix}.$$

The estimator  $\hat{\vartheta}_T$  is calculated from the log-likelihood function

$$\log \frac{dP_T^\vartheta}{dP_T^{(0,0)}}(X) = \vartheta^* V_T^0 - \frac{1}{2} \vartheta^* I_T^0 \vartheta, \quad \vartheta \in R_2$$

(see e.g. Liptser and Shirayev 1977). Here  $P_T^{(a,b)}$  is the measure on  $\mathbb{C}([-1, T])$  generated by the solution  $(X(t), t \in [-1, T])$  of (1.1).

The main purpose of this paper is to study local asymptotic properties of the family  $(P_T^\vartheta, \vartheta \in R_2)$  and then to draw conclusions for properties of the estimator  $\hat{\vartheta}_T$  when  $T \rightarrow \infty$ .

Since the log-likelihoods are quadratic in  $\vartheta$  for each  $T > 0$ , it is not surprising that the family  $(P_T^\vartheta), T > 0$ , is locally asymptotically quadratic (LAQ) at every  $\vartheta_0 \in R_2$ , see Section 2 (for the notion of LAQ see Le Cam and Yang, 1990; or Jeganathan 1995). Namely, choose  $\vartheta_0 = (a, b)^* \in R_2$  arbitrary but fixed and introduce  $\vartheta = \vartheta_0 + \varphi_T \mu$ , where  $\mu = (\alpha, \beta)^* \in R_2$ ,  $\varphi_T = \varphi_T(\vartheta_0)$  is a normalizing regular  $2 \times 2$  matrix with  $\varphi_T \rightarrow 0$  as  $T \rightarrow \infty$ . Then we get

$$\log \frac{dP_T^\vartheta}{dP_T^{\vartheta_0}}(X) = \mu^* V_T - \frac{1}{2} \mu^* I_T \mu, \quad (1.6)$$

where

$$V_T^* = \left( \int_0^T X(t) dW(t), \int_0^T X(t-1) dW(t) \right) \varphi_T \quad (1.7)$$

and

$$I_T = \varphi_T^* I_T^0 \varphi_T. \quad (1.8)$$

In view of (1.6), to prove LAQ at  $\vartheta_0$  one has to choose the matrices  $\varphi_T(\vartheta_0)$  in such a way that (a) the vectors  $(V_T, I_T)$  are bounded in probability as  $T \rightarrow \infty$ ; (b) if  $(V_{T_n}, I_{T_n})$  converges in distribution to a limit  $(V_\infty, I_\infty)$  for a subsequence  $\{T_n\} \rightarrow \infty$ , then

$$E \exp(\mu^* V_\infty - \frac{1}{2} \mu^* I_\infty \mu) = 1$$

for every  $\mu \in R_2$ ; (c) if  $I_{T_n}$  converges in distribution to a limit  $I_\infty$  for a subsequence  $\{T_n\} \rightarrow \infty$ , then  $I_\infty$  is almost surely positive definite. Recall also that the important special cases of LAQ are the local asymptotic mixed normality (LAMN) and the local asymptotic normality (LAN). LAMN at  $\vartheta_0$  means that  $(V_T, I_T)$  converges in distribution to  $(I_\infty^{1/2} Z, I_\infty)$  as  $T \rightarrow \infty$ , where the matrix  $I_\infty$  is almost surely positive definite and  $Z$  is a standard Gaussian vector independent of  $I_\infty$ . If, moreover,  $I_\infty$  is nonrandom, then we have LAN at  $\vartheta_0$ .

Note that condition (c) is important since otherwise we are not in a position even to

establish asymptotic properties of  $\hat{\vartheta}_T$  (cf. Dietz 1992). In general, (c) cannot be reached with matrices  $\varphi_T$  being diagonal. We construct  $\varphi_T$  as the product of two quadratic matrices  $\varphi_T^{(1)}$  and  $\varphi_T^{(2)}$ ,  $\varphi_T = \varphi_T^{(1)} \varphi_T^{(2)}$ , where  $\varphi_T^{(1)}$  converges to a nonsingular limit as  $T \rightarrow \infty$  (the dependence on  $T$  cannot be avoided in general) and  $\varphi_T^{(2)}$  is diagonal with elements tending to zero, in most case with different rates.

It is obvious from (1.7), (1.8) and (1.2) that the properties of the fundamental solution  $x_0(t)$  for  $t \rightarrow \infty$  very influence the limit properties of  $(V_T, I_T)$ . Recall that for Langevin's equation ( $b = 0$ ), it holds  $x_0(t) = e^{at}$ , the solution  $(X(t), t \geq 0)$  is the Ornstein–Uhlenbeck process and there are exactly three relevant cases in considering local asymptotic properties ( $a < 0$ ,  $a = 0$ ,  $a > 0$ ). In our case the picture turns out to be much more rich. To specify  $\varphi_T$  and to study the limit behaviour of  $(V_T, I_T)$  we have to distinguish eleven different cases for  $\vartheta_0$ . These cases will be introduced as follows. The behaviour of  $x_0(\cdot)$  is connected with the set  $\Lambda$  of (complex) solutions of the characteristic equation

$$\lambda - a - be^{-\lambda} = 0. \quad (1.9)$$

It is easy to see that the set  $\Lambda$  of solutions of (1.9) is countable infinite (if  $b \neq 0$ ) and that for every  $c \in R_1$  the set  $\Lambda_c := \{\lambda \in \Lambda \mid \operatorname{Re} \lambda \geq c\}$  is finite. In particular,  $v_0 := \max\{\operatorname{Re} \lambda \mid \lambda \in \Lambda\} < \infty$ . Define  $v_1 := \max\{\operatorname{Re} \lambda \mid \lambda \in \Lambda, \operatorname{Re} \lambda < v_0\}$  ( $\max \emptyset = -\infty$ ). One verifies easily that if  $\lambda \in \Lambda$  then  $\bar{\lambda} \in \Lambda$  and no other  $\mu \in \Lambda$  with  $\operatorname{Re} \mu = \operatorname{Re} \lambda$  exists. The equation (1.9) has at most two real solutions. If there exists a real solution  $v$  then the real part of every nonreal solution is strictly less than  $v$ . Consequently, the only possible real solutions are  $v_0$  (if there is exactly one) or  $v_0$  and  $v_1$  (if there are two).

We have  $v_0 \in \Lambda$  if and only if

$$b \geq v(a) := -e^{a-1}, \quad (1.10)$$

otherwise there exists a unique  $\lambda_0$  in  $\Lambda$  with  $\operatorname{Re} \lambda_0 = v_0$  and  $\xi_0 := \operatorname{Im} \lambda_0 > 0$ . Furthermore in this case it holds  $\xi_0 < \pi$ . Moreover, a second real solution exists, i.e.  $v_1 \in \Lambda$ , if and only if  $v(a) < b < 0$ .

For every  $\lambda$  in  $\Lambda$  denote by  $m(\lambda)$  the multiplicity of  $\lambda$  as a solution of (1.8). We have  $m(\lambda) = 1$  for all  $\lambda \in \Lambda$  except  $\lambda = a - 1$ , which belongs to  $\Lambda$  if and only if  $b = v(a)$  and which has multiplicity two, if  $b = v(a)$ .

Additional information about the solutions of the equation (1.9) can be found in Hayes (1950).

The following lemma is crucial for this note. It is based on the inverse Laplace transform and Cauchy's residual theorem and it can be found in a slightly other form in Myschkis (1972), see also Hale and Verduyn Lunel (1993). The proof will be sketched in the appendix.

**Lemma 1.1** *For all  $c < v_0$  the fundamental solution  $x_0(\cdot)$  of (1.3) can be represented in the form*

$$x_0(t) = \psi_0(t)e^{v_0 t} + \sum_{\substack{\lambda_k \in \Lambda_c \\ \operatorname{Re} \lambda_k < v_0}} c_k e^{\lambda_k t} + o(e^{\gamma t}) \quad \text{for } t \rightarrow \infty, \quad (1.11)$$

where  $\gamma < c$  and  $c_k$  are some constants. Here  $\psi_0(t)$  equals

$$\psi_0(t) = \begin{cases} \frac{1}{v_0 - a + 1} & \text{if } v_0 \in \Lambda, m(v_0) = 1, \\ 2t + \frac{2}{3} & \text{if } v_0 \in \Lambda, m(v_0) = 2, \\ A_0 \cos(\xi_0 t) + B_0 \sin(\xi_0 t) & \text{if } v_0 \notin \Lambda, \end{cases}$$

$$A_0 = \frac{2(v_0 - a + 1)}{(v_0 - a + 1)^2 + \xi_0^2}, \quad B_0 = \frac{2\xi_0}{(v_0 - a + 1)^2 + \xi_0^2}.$$

**Remarks:**

- 1) Note that the three cases for  $\psi_0$  correspond to  $b > v(a)$ ,  $b = v(a)$  and  $b < v(a)$ , respectively.
- 2) Recall that for Langevin's equation, i.e.  $b = 0$ , we have  $b > v(a)$  for every  $a \in R_1$ . In this case it holds  $\Lambda = \{a\}$  and therefore  $v_0 = a$  and  $x_0(t) = e^{at}$ .
- 3) If  $v_0 \in \Lambda$ ,  $m(v_0) = 1$  (and  $b \neq 0$  to avoid the case from the previous remark), then for our purposes it is necessary to separate a further term from the sum in (1.11). We get

$$x_0(t) = \frac{1}{v_0 - a + 1} e^{v_0 t} + \psi_1(t) e^{v_1 t} + o(e^{\gamma t}) \quad \text{for } t \rightarrow \infty, \quad (1.12)$$

where  $\gamma < v_1$ ,

$$\psi_1(t) = \begin{cases} \frac{1}{v_1 - a + 1} & \text{if } v_1 \in \Lambda, \\ A_1 \cos(\xi_1 t) + A_2 \sin(\xi_1 t) & \text{if } v_1 \notin \Lambda, \end{cases}$$

$$A_1 = \frac{2(v_1 - a + 1)}{(v_1 - a + 1)^2 + \xi_1^2}, \quad B_1 = \frac{2\xi_1}{(v_1 - a + 1)^2 + \xi_1^2}.$$

Here  $\xi_1$  denotes the uniquely determined positive number such that  $\lambda_1 = v_1 + i\xi_1 \in \Lambda$ . (We note that  $\xi_1 \in (\pi, 2\pi)$  in this case.)

The proof follows the line of the proof of Lemma 1.1 (see the Appendix) in an obvious way.

As it was mentioned above, the limit properties of  $(V_T, I_T)$  are different in eleven cases. Table 1 represents these cases. The first column describes these cases in terms of  $v_0$  and  $v_1$ , and the relations (1.11) and (1.12) make clear a connection between our classification and asymptotic properties of  $x_0(\cdot)$ . The second column characterizes the cases in terms of  $a$  and  $b$ . The third column is just a notation for these cases which will be used in the rest of the paper. The functions  $u(a)$ ,  $a < 1$ , and  $w(a)$ ,  $a \in R_1$ ,

are defined as follows: introduce a parametric curve  $(a(\xi), b(\xi))$ ,  $\xi > 0$ ,  $\xi \neq \pi, 2\pi, \dots$ , in  $R_2$  by

$$a(\xi) = \xi \cot \xi, \quad b(\xi) = -\xi / \sin \xi,$$

then  $b = u(a)$  and  $b = w(a)$  are the branches of this curve corresponding to  $\xi \in (0, \pi)$  and  $\xi \in (\pi, 2\pi)$  respectively, see also Figure 1.

$v_0 < 0$				$a < 1, \ u(a) < b < -a$	N	
$v_0 = 0$	$v_0 \in \Lambda$	$m(v_0) = 1$		$a < 1, \ b = -a$	Q1	
		$m(v_0) = 2$		$a = 1, \ b = -a$	Q2	
	$v_0 \notin \Lambda$		$a < 1, \ b = u(a)$		Q3	
$v_0 > 0$	$v_0 \in \Lambda$	$m(v_0) = 1$	$v_1 < 0$		$-a < b < w(a)$	M1
			$v_1 = 0$	$v_1 \in \Lambda$	$a > 1, \ b = -a$	Q4
				$v_1 \notin \Lambda$	$b = w(a)$	Q5
			$v_1 > 0$	$v_1 \in \Lambda$	$a > 1, \ v(a) < b < -a$	M2
				$v_1 \notin \Lambda$	$b > w(a)$	P1
		$m(v_0) = 2$		$a > 1, \ b = v(a)$		M3
	$v_0 \notin \Lambda$		$a < 1, b < u(a)$ or $a \geq 1, b < v(a)$		P2	

Table 1: The eleven cases in terms of  $(v_0, v_1)$  and  $(a, b)$

In the following we want to give a first impression what happens in the eleven cases. The first subdivision on the left-hand side of Table 1 follows the Ornstein–Uhlenbeck case (where  $v_0 = a$ ):  $v_0 < 0$ ,  $v_0 = 0$  and  $v_0 > 0$ .

The first case  $v_0 < 0$  holds if and only if there exists a stationary solution of (1.1). This solution is Gaussian and uniquely determined (see K  chler and Mensch 1991). In this case the statistical properties of our model are classical: the LAN property holds. The form of  $\psi_0(\cdot)$  does not influence the asymptotic properties of  $V_T$  and  $I_T$  if  $v_0 < 0$ . But it does if  $v_0 = 0$  or  $v_0 > 0$ .

If  $v_0 = 0$ , the underlying experiment is only LAQ (as if  $a = 0$  in the Ornstein–Uhlenbeck case) in all three cases Q1–Q3. But the normalizing matrix  $\varphi_T$  and the corresponding limit experiment now principally depend on the form of  $\psi_0(\cdot)$ , represented by three different expressions in Lemma 1.1.

Now let us consider the case  $v_0 > 0$ . The form of  $\psi_0(\cdot)$  is essential again for our purposes. If  $v_0 \notin \Lambda$ , then we get a periodic behaviour of  $(V_T, I_T)$  in a certain sense. We call it the periodically locally asymptotically mixed normal (PLAMN) property to emphasize that the cluster points of  $(V_T, I_T)$  have the same structure as in the LAMN case but  $(V_T, I_T)$  converges in distribution if  $T$  runs to infinity through a grid with a fixed step. If  $v_0 \in \Lambda$  and  $m(v_0) = 2$ , then the model is locally asymptotically mixed normal. This is the only case where the matrix  $\varphi_T^{(1)}$  has to be chosen dependent on  $T$ . If  $v_0 \in \Lambda$  and  $m(v_0) = 1$ , we have to take into consideration also the second term on the right-hand side of (1.12) and obtain five different cases. Indeed, the limit behaviour depends on the sign of  $v_1$  (i.e.  $v_1 < 0$ ,  $v_1 = 0$  or  $v_1 > 0$ ) and, if  $v_1 \geq 0$ , on the fact if  $v_1 \in \Lambda$  or not.

If  $v_1 < 0$ , or  $v_1 > 0$  and  $v_1 \in \Lambda$ , the underlying experiment is locally asymptotically mixed normal. But if  $v_1 > 0$  and  $v_1 \notin \Lambda$ , a periodic behaviour of  $(V_T, I_T)$  occurs again and we have PLAMN.

If  $v_1 = 0$ , the model is locally asymptotically quadratic, whether  $v_1 \in \Lambda$  or not. Both cases have similar limit experiments.

Finally, we note that LAMN (or PLAMN) fails only if  $\vartheta_0$  belongs to the lines  $b = -a$ ,  $b = u(a)$ , or  $b = w(a)$ .

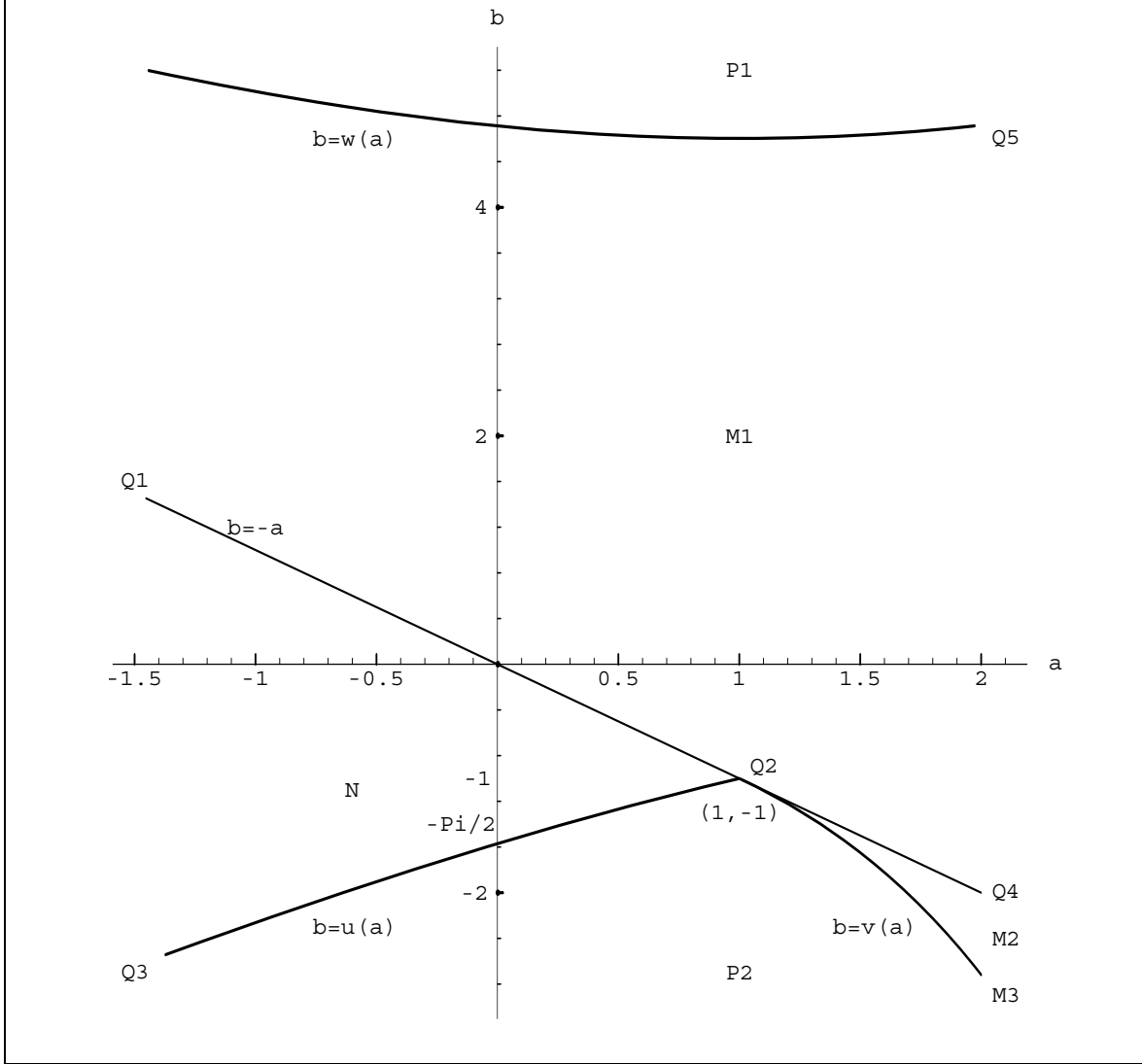


Figure 1: The different cases for  $\vartheta_0 = (a, b)^*$

## 2 Local asymptotic properties

In the preceding section we have introduced a series of cases for which the fundamental solution  $x_0(\cdot)$  and  $(V_T, I_T)$  have different asymptotic properties. Here we shall study the asymptotic properties of  $(V_T, I_T)$  as  $T \rightarrow \infty$  in more detail. The proofs are given

in Section 4.

In the sequel the symbols  $\xrightarrow{P}$  and  $\xrightarrow{d}$  denote convergence in probability and in distribution respectively. We shall use the symbol  $\xrightarrow{\mathcal{L}}$  to denote the convergence in distribution in the space  $\mathbb{C}^d([0, 1])$  of continuous functions on  $[0, 1]$  with values in  $R^d$ . Sometimes we shall use the abbreviated notation  $\int_0^1 \widetilde{W}_1 d\widetilde{W}_1$  instead of  $\int_0^1 \widetilde{W}_1(t) d\widetilde{W}_1(t)$  or  $\int_0^1 XY dt$  instead of  $\int_0^1 X(t)Y(t) dt$  etc. The concrete meaning will be clear from the context.

Some details are represented in the next table. In the first column all the eleven cases are listed in the order in which they will be considered. The next three columns describe the choice of  $\varphi_T$ . Recall that

$$\varphi_T = \varphi_T^{(1)} \varphi_T^{(2)} \quad \text{and} \quad \varphi_T^{(2)} = \begin{pmatrix} \varphi_{11}(T) & 0 \\ 0 & \varphi_{22}(T) \end{pmatrix}. \quad (2.1)$$

It turns out that we can choose  $\varphi_T^{(1)}$  so that it does not depend on  $T$  in all the cases except M3; moreover,  $\varphi_T^{(1)}$  equals

$$\text{either} \quad \mathcal{I}_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \mathcal{J}_2 := \begin{pmatrix} 1 & 1 \\ 0 & -e^{v_0} \end{pmatrix}$$

(of course, the value of  $v_0$  depends on  $\vartheta_0$ ).

It will be proved in Propositions 2.1, 2.2–2.4, 2.7–2.11 that in Cases N, M1–M3, Q1–Q5, under this choice of  $\varphi_T$ , we have  $(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty)$ . It is implicitly assumed that

$$E \exp(\mu^* V_\infty - \frac{1}{2} \mu^* I_\infty \mu) = 1$$

for every  $\mu \in R_2$  and  $I_\infty$  is a non-singular matrix; the proof is either trivial or routine. In the fifth column of Table 2 we describe the matrix  $I_\infty$  and the type of convergence of  $I_T$  to  $I_\infty$  in a symbolic manner. Namely, the elements of the matrix in this column have the following meaning. The symbol ‘0’ means that the corresponding element of  $I_\infty$  is 0. The symbol ‘c’ means that the corresponding element of  $I_\infty$  is a (nonzero) constant. In both cases, the corresponding element of  $I_T$  converges to this constant in probability. The symbol ‘p’ means that the corresponding element of  $I_\infty$  is random but there is still the convergence in probability of the corresponding element of  $I_T$  to that of  $I_\infty$ . Finally, the symbol ‘d’ means that we have only the convergence in distribution of the corresponding element of  $I_T$  to that of  $I_\infty$  but not the convergence in probability. In Cases P1 and P2 studied in Propositions 2.5 and 2.6, we have a periodic behaviour of  $(V_T, I_T)$  in a certain sense. We use there the symbol ‘p\*’ to indicate that we have the convergence in probability of the corresponding elements of  $I_T$  to a random limit but only when  $T$  runs to infinity through certain grids.

The last column of the table indicates the number of the proposition in which the corresponding case is considered.



Case	$\varphi_T^{(1)}$	$\varphi_{11}(T)$	$\varphi_{22}(T)$	$I_\infty$	Prop.
N	$\mathcal{I}_2$	$T^{-1/2}$	$T^{-1/2}$	$\begin{pmatrix} c & c \\ c & c \end{pmatrix}$	2.1
M1	$\mathcal{J}_2$	$e^{-v_0 T}$	$T^{-1/2}$	$\begin{pmatrix} p & 0 \\ 0 & c \end{pmatrix}$	2.2
M2	$\mathcal{J}_2$	$e^{-v_0 T}$	$e^{-v_1 T}$	$\begin{pmatrix} p & p \\ p & p \end{pmatrix}$	2.3
M3	$\begin{pmatrix} 1 & 1 \\ 0 & -(1 + T^{-1})e^{v_0} \end{pmatrix}$	$T^{-1}e^{-v_0 T}$	$Te^{-v_0 T}$	$\begin{pmatrix} p & p \\ p & p \end{pmatrix}$	2.4
P1	$\mathcal{J}_2$	$e^{-v_0 T}$	$e^{-v_1 T}$	$\begin{pmatrix} p & p^* \\ p^* & p^* \end{pmatrix}$	2.5
P2	$\mathcal{I}_2$	$e^{-v_0 T}$	$e^{-v_0 T}$	$\begin{pmatrix} p^* & p^* \\ p^* & p^* \end{pmatrix}$	2.6
Q1	$\mathcal{J}_2$	$T^{-1}$	$T^{-1/2}$	$\begin{pmatrix} d & 0 \\ 0 & c \end{pmatrix}$	2.7
Q2	$\mathcal{J}_2$	$T^{-2}$	$T^{-1}$	$\begin{pmatrix} d & d \\ d & d \end{pmatrix}$	2.8
Q3	$\mathcal{I}_2$	$T^{-1}$	$T^{-1}$	$\begin{pmatrix} d & d \\ d & d \end{pmatrix}$	2.9
Q4	$\mathcal{J}_2$	$e^{-v_0 T}$	$T^{-1}$	$\begin{pmatrix} p & 0 \\ 0 & d \end{pmatrix}$	2.10
Q5	$\mathcal{J}_2$	$e^{-v_0 T}$	$T^{-1}$	$\begin{pmatrix} p & 0 \\ 0 & d \end{pmatrix}$	2.11

Table 2: The choice of  $\varphi_T$  and a description of convergence to  $I_\infty$

In the following we shall treat every case mentioned above in a separate proposition. Recall that  $V_T$  and  $I_T$  are given by (1.7) and (1.8). The process  $X(\cdot)$  is defined by (1.2) for some fixed  $a$  and  $b$  and the matrices  $\varphi_T$  are constructed in (2.1). For every proposition below the parameter  $\vartheta_0 = (a, b)^*$  is assumed to belong to the set described by Table 1 in accordance with the case under consideration. The definitions of  $\varphi_T^{(1)}$ ,  $\varphi_{11}(T)$  and  $\varphi_{22}(T)$  have to be taken from Table 2. Unless otherwise specified, all limits are taken as  $T \rightarrow \infty$ .

Let us start with the most simple case  $v_0 < 0$ . This is the only case where  $\int_0^\infty x_0^2(t) dt < \infty$  and a stationary solution of (1.1) exists.

**Proposition 2.1** *In Case N the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically normal at every  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where

$$I_\infty = \begin{pmatrix} \int_0^\infty x_0^2(t) dt & \int_0^\infty x_0(t)x_0(t+1) dt \\ \int_0^\infty x_0(t)x_0(t+1) dt & \int_0^\infty x_0^2(t) dt \end{pmatrix}$$

and  $V_\infty \sim \mathcal{N}(0, I_\infty)$ .

Now let us treat Cases M1–M3.

**Proposition 2.2** *In Case M1 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically mixed normal at every  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty) \stackrel{d}{=} (I_\infty^{1/2}Z, I_\infty)$  and  $Z$  is an independent of  $I_\infty$  and  $\mathcal{N}(0, \mathcal{I}_2)$ -distributed vector. The matrix  $I_\infty$  is given by

$$I_\infty = \begin{pmatrix} \frac{U_0^2}{2v_0(v_0 - a + 1)^2} & 0 \\ 0 & \int_0^\infty (x_0(t) - e^{v_0}x_0(t-1))^2 dt \end{pmatrix},$$

where

$$U_0 = X_0(0) + b \int_{-1}^0 e^{-v_0(s+1)} X_0(s) ds + \int_0^\infty e^{-v_0 s} dW(s).$$

**Proposition 2.3** *In Case M2 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically mixed normal at every  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty) \stackrel{d}{=} (I_\infty^{1/2}Z, I_\infty)$  and  $Z$  is an independent of  $I_\infty$  and  $\mathcal{N}(0, \mathcal{I}_2)$ -distributed vector. The matrix  $I_\infty$  is given by

$$I_\infty = \begin{pmatrix} \frac{U_0^2}{2v_0(v_0 - a + 1)^2} & \frac{U_0 U_1 (e^{v_0 - v_1} - 1)}{(v_0 + v_1)(v_0 - a + 1)(a - v_1 - 1)} \\ \frac{U_0 U_1 (e^{v_0 - v_1} - 1)}{(v_0 + v_1)(v_0 - a + 1)(a - v_1 - 1)} & \frac{U_1^2 (e^{v_0 - v_1} - 1)^2}{2v_1(a - v_1 - 1)^2} \end{pmatrix},$$

where  $U_0$  is defined in Proposition 2.2 and

$$U_1 = X_0(0) + b \int_{-1}^0 e^{-v_1(s+1)} X_0(s) ds + \int_0^\infty e^{-v_1 s} dW(s).$$

**Proposition 2.4** *In Case M3 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically mixed normal at  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty) \stackrel{d}{=} (I_\infty^{1/2}Z, I_\infty)$  and  $Z$  is an independent of  $I_\infty$  and  $\mathcal{N}(0, \mathcal{I}_2)$ -distributed vector. The matrix  $I_\infty$  is given by

$$I_\infty = \begin{pmatrix} \frac{2U_0^2}{v_0} & \frac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0} + \frac{U_0^2}{v_0^2} \\ \frac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0} + \frac{U_0^2}{v_0^2} & \frac{(\frac{4}{3}U_0 + 2U_2)^2}{2v_0} + \frac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0^2} + \frac{U_0^2}{v_0^3} \end{pmatrix},$$

where  $U_0$  is defined in Proposition 2.2 and

$$U_2 = b \int_{-1}^0 (s+1)e^{-v_0(s+1)} X_0(s) ds + \int_0^\infty s e^{-v_0 s} dW(s).$$

The next two propositions treat Cases P1 and P2. Let us recall that if  $v_i \notin \Lambda$ , then  $\xi_i$  denotes the positive imaginary part of  $\lambda \in \Lambda$  with  $\operatorname{Re} \lambda = v_i$ ,  $i = 0, 1$ .

**Proposition 2.5** *In Case P1 the family  $(P^\vartheta, \vartheta \in R_2)$  is “periodically locally asymptotically mixed normal” at every  $\vartheta_0$  in the following sense: for  $T_n = u + n\Delta$ , where  $u \in [0, \Delta)$  is fixed,  $\Delta = 2\pi/\xi_1$ ,  $n \geq 0$ ,*

$$(V_{T_n}, I_{T_n}) \xrightarrow{d} (V_\infty(u), I_\infty(u)), \quad n \rightarrow \infty,$$

where  $(V_\infty(u), I_\infty(u)) \stackrel{d}{=} (I_\infty^{1/2}(u)Z, I_\infty(u))$  and  $Z$  is an independent of  $I_\infty(u)$  and  $\mathcal{N}(0, \mathcal{I}_2)$ -distributed random vector. The matrix  $I_\infty(u)$  is given by

$$I_\infty(u) = \begin{pmatrix} \frac{U_0^2}{2v_0(v_0 - a + 1)^2} & \frac{U_0}{v_0 - a + 1} \int_0^\infty e^{-(v_0+v_1)t} U(u-t) dt \\ \frac{U_0}{v_0 - a + 1} \int_0^\infty e^{-(v_0+v_1)t} U(u-t) dt & \int_0^\infty e^{-2v_1 t} U^2(u-t) dt \end{pmatrix}.$$

Here  $U_0$  is defined in Proposition 2.2,

$$U(t) = X_0(0)\phi(t) + b \int_{-1}^0 \phi(t-s-1)e^{-v_1(s+1)}X_0(s)ds + \int_0^\infty \phi(t-s)e^{-v_1 s}dW(s),$$

$$\phi(t) = A \cos(\xi_1 t) + B \sin(\xi_1 t)$$

and

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} - e^{v_0-v_1} \begin{pmatrix} \cos \xi_1 & -\sin \xi_1 \\ \sin \xi_1 & \cos \xi_1 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix},$$

where

$$A_1 = \frac{2(v_1 - a + 1)}{(v_1 - a + 1)^2 + \xi_1^2}, \quad B_1 = \frac{2\xi_1}{(v_1 - a + 1)^2 + \xi_1^2}.$$

**Proposition 2.6** *In Case P2 the family  $(P^\vartheta, \vartheta \in R_2)$  is “periodically locally asymptotically mixed normal” at every  $\vartheta_0$  in the following sense: for  $T_n = u + n\Delta$ , where  $u \in [0, \Delta)$  is fixed,  $\Delta = \pi/\xi_0$ ,  $n \geq 0$ ,*

$$(V_{T_n}, I_{T_n}) \xrightarrow{d} (V_\infty(u), I_\infty(u)), \quad n \rightarrow \infty,$$

where  $(V_\infty(u), I_\infty(u)) \stackrel{d}{=} (I_\infty^{1/2}(u)Z, I_\infty(u))$  and  $Z$  is an independent of  $I_\infty(u)$  and  $\mathcal{N}(0, \mathcal{I}_2)$ -distributed random vector. The matrix  $I_\infty(u)$  is given by

$$I_\infty(u) = \begin{pmatrix} \int_0^\infty e^{-2v_0 t} U_0^2(u-t) dt & \int_0^\infty e^{-2v_0 t} U_0(u-t)U_2(u-t) dt \\ \int_0^\infty e^{-2v_0 t} U_0(u-t)U_2(u-t) dt & \int_0^\infty e^{-2v_0 t} U_2^2(u-t) dt \end{pmatrix},$$

where

$$U_i(t) = X_0(0)\phi_i(t) + b \int_{-1}^0 \phi_i(t-s-1)e^{-v_0(s+1)}X_0(s)ds + \int_0^\infty \phi_i(t-s)e^{-v_0 s}dW(s),$$

$$\begin{aligned}\phi_i(t) &= A_i \cos(\xi_0 t) + B_i \sin(\xi_0 t), \quad i = 0, 2, \\ A_0 &= \frac{2(v_0 - a + 1)}{(v_0 - a + 1)^2 + \xi_0^2}, \quad B_0 = \frac{2\xi_0}{(v_0 - a + 1)^2 + \xi_0^2}, \\ \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} &= e^{-v_0} \begin{pmatrix} \cos \xi_0 & -\sin \xi_0 \\ \sin \xi_0 & \cos \xi_0 \end{pmatrix} \begin{pmatrix} A_0 \\ B_0 \end{pmatrix}.\end{aligned}$$

So far we have treated all the cases for which LAN, LAMN or PLAMN holds. There remain five cases, where LAQ is valid.

**Proposition 2.7** *In Case Q1 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically quadratic at every  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty)$  is given by

$$V_\infty = \left( \frac{1}{1-a} \int_0^1 \widetilde{W}(t) d\widetilde{W}(t), Z \right)$$

and

$$I_\infty = \begin{pmatrix} \frac{1}{(1-a)^2} \int_0^1 \widetilde{W}^2(t) dt & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

Here  $\sigma^2 = \int_0^\infty (x_0(t) - x_0(t-1))^2 dt$ ,  $(\widetilde{W}(t), t \in [0, 1])$  denotes a standard Wiener process and  $Z$  is a  $\mathcal{N}(0, \sigma^2)$ -distributed random variable independent of  $\widetilde{W}(\cdot)$ .

**Proposition 2.8** *In Case Q2 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically quadratic at  $\vartheta_0 = (1, -1)^*$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty)$  is given by

$$\begin{aligned}V_\infty &= 2 \left( \int_0^1 \widetilde{X}(t) d\widetilde{W}(t), \int_0^1 \widetilde{W}(t) d\widetilde{W}(t) \right), \\ I_\infty &= 4 \begin{pmatrix} \int_0^1 \widetilde{X}^2(t) dt & \int_0^1 \widetilde{X}(t) \widetilde{W}(t) dt \\ \int_0^1 \widetilde{X}(t) \widetilde{W}(t) dt & \int_0^1 \widetilde{W}^2(t) dt \end{pmatrix}.\end{aligned}$$

Here  $(\widetilde{W}(t), t \in [0, 1])$  is a standard Wiener process and  $\widetilde{X}(t) = \int_0^t \widetilde{W}(s) ds$ .

**Proposition 2.9** *In Case Q3 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically quadratic at every  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where

$$V_\infty = \frac{1}{2} \begin{pmatrix} A_0 \int_0^1 \widetilde{W}_1 d\widetilde{W}_1 + A_0 \int_0^1 \widetilde{W}_2 d\widetilde{W}_2 + B_0 \int_0^1 \widetilde{W}_1 d\widetilde{W}_2 - B_0 \int_0^1 \widetilde{W}_2 d\widetilde{W}_1 \\ A_2 \int_0^1 \widetilde{W}_1 d\widetilde{W}_1 + A_2 \int_0^1 \widetilde{W}_2 d\widetilde{W}_2 + B_2 \int_0^1 \widetilde{W}_1 d\widetilde{W}_2 - B_2 \int_0^1 \widetilde{W}_2 d\widetilde{W}_1 \end{pmatrix}$$

and

$$I_\infty = \frac{A_0^2 + B_0^2}{4} \begin{pmatrix} 1 & \cos \xi_0 \\ \cos \xi_0 & 1 \end{pmatrix} \int_0^1 (\widetilde{W}_1^2 + \widetilde{W}_2^2) dt.$$

Here  $(\widetilde{W}_i(t), t \in [0, 1])$ ,  $i = 1, 2$ , are two independent standard Wiener processes and

$$A_0 = \frac{2(1-a)}{(1-a)^2 + \xi_0^2}, \quad B_0 = \frac{2\xi_0}{(1-a)^2 + \xi_0^2},$$

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} \cos \xi_0 & -\sin \xi_0 \\ \sin \xi_0 & \cos \xi_0 \end{pmatrix} \begin{pmatrix} A_0 \\ B_0 \end{pmatrix}.$$

**Proposition 2.10** *In Case Q4 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically quadratic at every  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty)$  is given by

$$V_\infty = \left( \frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{e^{v_0} - 1}{a - 1} \int_0^1 \widetilde{W}(t) d\widetilde{W}(t) \right)^*$$

and

$$I_\infty = \begin{pmatrix} \frac{U_0^2}{2v_0(v_0 - a + 1)^2} & 0 \\ 0 & \frac{(e^{v_0} - 1)^2}{(a - 1)^2} \int_0^1 \widetilde{W}^2(t) dt \end{pmatrix}.$$

Here  $U_0$  is the same as in Proposition 2.2 above,  $Z$  and  $(\widetilde{W}(t), t \in [0, 1])$  are a standard normal distributed random variable and a standard Wiener process respectively, and  $U_0, Z$  and  $\widetilde{W}(\cdot)$  are independent.

**Proposition 2.11** *In Case Q5 the family  $(P^\vartheta, \vartheta \in R_2)$  is locally asymptotically quadratic at  $\vartheta_0$ :*

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty),$$

where  $(V_\infty, I_\infty)$  is given by

$$V_\infty = \begin{pmatrix} \frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)} \\ \frac{1}{2} \left( A \int_0^1 \widetilde{W}_1 d\widetilde{W}_1 + A \int_0^1 \widetilde{W}_2 d\widetilde{W}_2 + B \int_0^1 \widetilde{W}_1 d\widetilde{W}_2 - B \int_0^1 \widetilde{W}_2 d\widetilde{W}_1 \right) \end{pmatrix}$$

and

$$I_\infty = \begin{pmatrix} \frac{U_0^2}{2v_0(v_0 - a + 1)^2} & 0 \\ 0 & \frac{1}{4}(A^2 + B^2) \int_0^1 (\widetilde{W}_1^2 + \widetilde{W}_2^2) dt \end{pmatrix}.$$

Here  $U_0$  is the same as in Proposition 2.2 above,  $Z$  and  $(\widetilde{W}_i(t), t \in [0, 1])$ ,  $i = 1, 2$ , are a standard normal distributed random variable and standard Wiener processes respectively, and  $U_0, Z, \widetilde{W}_1(\cdot), \widetilde{W}_2(\cdot)$  are independent,

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} - e^{v_0} \begin{pmatrix} \cos \xi_1 & -\sin \xi_1 \\ \sin \xi_1 & \cos \xi_1 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix},$$

where

$$A_1 = \frac{2(1 - a)}{(1 - a)^2 + \xi_1^2}, \quad B_1 = \frac{2\xi_1}{(1 - a)^2 + \xi_1^2}.$$

### 3 Asymptotic Properties of the ML-Estimator

Assume we observe  $(X(t), t \leq T)$  continuously, where  $X(t)$  is a solution (1.1) and the parameters  $a$  and  $b$  are unknown. The maximum likelihood estimator  $\hat{\vartheta}_T$  of the true parameter  $\vartheta_0 = (a_0, b_0)^*$  is given by

$$\hat{\vartheta}_T = \arg \max_{\vartheta \in R_2} \ell_T^0(\vartheta) = (I_T^0)^{-1} V_T^0,$$

where

$$\ell_T^0(\vartheta) = \vartheta^* V_T^0 - \frac{1}{2} \vartheta^* I_T^0 \vartheta, \quad \vartheta \in R_2,$$

$$V_T^0 = \left( \int_0^T X(t) dX(t), \int_0^T X(t-1) dX(t) \right)^*$$

and

$$I_T^0 = \begin{pmatrix} \int_0^T X^2(t) dt & \int_0^T X(t)X(t-1) dt \\ \int_0^T X(t)X(t-1) dt & \int_0^T X^2(t-1) dt \end{pmatrix}.$$

Choose an arbitrary nonsingular  $2 \times 2$ -matrix  $\varphi_T$  and introduce a new parameter  $\mu = (\alpha, \beta)^* \in R_2$  by

$$\vartheta = \vartheta_0 + \varphi_T \mu.$$

Then

$$\hat{\vartheta}_T = \vartheta_0 + \varphi_T \hat{\mu}_T,$$

where  $\hat{\mu}_T$  is defined by

$$\hat{\mu}_T = \arg \max_{\mu \in R_2} \ell_T(\mu) = I_T^{-1} V_T$$

with

$$\ell_T(\mu) = \mu^* V_T - \frac{1}{2} \mu^* I_T \mu,$$

$$V_T = \varphi_T^* \left( \int_0^T X(t) dW(t), \int_0^T X(t-1) dW(t) \right)^*$$

and

$$I_T = \varphi_T^* I_T^0 \varphi_T.$$

From Section 2 we know that under appropriate choice of  $\varphi_T$  we have <sup>2)</sup>

$$(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty) \quad (3.1)$$

or

$$(V_{u+n\Delta}, I_{u+n\Delta}) \xrightarrow{d} (V_\infty(u), I_\infty(u)) \quad (3.2)$$

with  $\det I_\infty \neq 0$  and  $\det I_\infty(u) \neq 0$  for every  $u \in [0, \Delta)$  respectively. Consequently we get

$$\varphi_T^{-1}(\hat{\vartheta}_T - \vartheta_0) = \hat{\mu}_T \xrightarrow{d} I_\infty^{-1} V_\infty$$

or

$$\varphi_{u+n\Delta}^{-1}(\hat{\vartheta}_{u+n\Delta} - \vartheta_0) = \hat{\mu}_{u+n\Delta} \xrightarrow{d} I_\infty^{-1}(u) V_\infty(u)$$

for every  $u \in [0, \Delta)$  respectively.

Thus we can draw conclusions concerning the asymptotic behaviour of  $\hat{\vartheta}_T$  for  $T \rightarrow \infty$ . But some more properties follow from (3.1) and (3.2) by standard arguments. Indeed, if the LAMN property holds (e.g. in Cases N, M1–M3) then we have the local asymptotic minimax bound for an arbitrary estimator  $\tilde{\vartheta}_T$

$$\begin{aligned} \lim_{r \rightarrow \infty} \liminf_{T \rightarrow \infty} \sup_{\|\varphi_T^{-1}(\tilde{\vartheta}_T - \vartheta_0)\| \leq r} E_{\vartheta} w\{\varphi_T^{-1}(\tilde{\vartheta}_T - \vartheta_0)\} &\geq E w\{(I_\infty)^{-1} V_\infty\} \\ &= E w\{(I_\infty)^{-1/2} Z\}, \end{aligned} \quad (3.3)$$

where  $Z$  is a  $\mathcal{N}(0, \mathcal{I}_2)$ -vector independent of  $I_\infty$  and  $w : R^2 \rightarrow [0, \infty)$  is a bowl shape loss function. The maximum likelihood estimator  $\hat{\vartheta}_T$  attains this bound, at least for bounded  $w$ . Moreover, the estimator  $\hat{\vartheta}_T$  is asymptotically efficient in the convolution theorem sense. (See e.g. Le Cam and Yang 1990; Jeganathan 1995.)

In other cases, e.g. if only the LAQ property holds, it follows that there exists a lower asymptotic minimax bound but may be of different form (see Shiryaev and Spokoiny 1997; Greenwood and Wefelmeyer 1993). This bound need not be attainable. It is known that the maximum likelihood estimator is asymptotically generalized Bayesian with respect to the uniform distribution on  $R_2$  (Shiryaev and Spokoiny 1997).

For some class of estimators  $\tilde{\vartheta}_T$  satisfying certain conditions of regularity, e.g. that the limit distribution of the randomly normed deviation

$$I_T \varphi_T^{-1}(\tilde{\vartheta}_T - \vartheta_0)$$

exists and is unbiased, the covariance matrix of this limit distribution is bounded from below by the corresponding covariance matrix for the maximum likelihood estimator  $\hat{\vartheta}_T$  which is equal to  $E I_\infty$ , see Gushchin (1995).

---

<sup>2)</sup> For the notation see Section 2 above.

We have seen that the maximum likelihood estimator after a certain matrix-normalization, converges in distribution to some limit. In Cases N, P2 and Q3 we have that  $\varphi_T$  is equal  $\varphi_{22}(T)\mathcal{I}_2$ , and thus the normalization by the number  $\varphi_{22}^{-1}(T)$  yields the same limit distribution.

In all other cases  $\varphi_T^{(1)}$  is an upper triangular matrix and  $\varphi_{11}(T) = o(\varphi_{22}(T))$ . This reflects some singularities in the local structure of our model, which have not been mentioned so far.

So we shall assume in the rest of this section that the true value  $\vartheta_0 = (a, b)^*$  of the parameter corresponds to one of Cases M1–M3, P1, Q1, Q2, Q4, or Q5. First, we note that the normalization of  $\hat{\vartheta}_T - \vartheta_0$  by the scalar  $\varphi_{22}^{-1}(T)$  leads to a nontrivial limit distribution which is concentrated on a straight line <sup>3)</sup>. Indeed, we get

$$\begin{aligned} \varphi_{22}^{-1}(T)(\hat{\vartheta}_T - \vartheta_0) &= \varphi_{22}^{-1}(T)\varphi_T\hat{\mu}_T = \varphi_{22}^{-1}(T)\varphi_T^{(1)}\varphi_T^{(2)}\hat{\mu}_T \\ &= \varphi_T^{(1)} \begin{pmatrix} \varphi_{22}^{-1}(T)\varphi_{11}(T) & 0 \\ 0 & 1 \end{pmatrix} \hat{\mu}_T \\ &\xrightarrow{d} \begin{pmatrix} 1 & 1 \\ 0 & -e^{v_0} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} I_\infty^{-1} V_\infty \\ &= \begin{pmatrix} 0 & 1 \\ 0 & -e^{v_0} \end{pmatrix} I_\infty^{-1} V_\infty. \end{aligned} \tag{3.4}$$

In particular, the rate of convergence of  $\hat{\vartheta}_T$  to  $\vartheta_0$  is  $\varphi_{22}^{-1}(T)$  (as in Cases N, P2, and Q3), and the distance between  $\hat{\vartheta}_T$  and the straight line  $\Theta_0 := \{(a', b') : (a' - a) + e^{-v_0}(b' - b) = 0\}$  is of smaller order than  $\varphi_{22}(T)$ .

In this connection it is of some interest to see what happens if one of the parameters or a linear combination of them is known. Here we shall concentrate on the maximum likelihood estimators. The corresponding arguments concerning local asymptotic properties are similar and omitted.

Assume the parameter  $\vartheta$  belongs to a straight line  $\Theta'$  which meets  $\vartheta_0$ . The limit behaviour of the maximum likelihood estimator

$$\hat{\vartheta}'_T = \arg \max_{\vartheta \in \Theta'} \ell_T^0(\vartheta)$$

is essentially different in the following two cases.

Case 1:  $\Theta' \neq \Theta_0$ .

Case 2:  $\Theta' = \Theta_0$ .

Denote by  $\mathfrak{M}_T$  the image of  $\Theta'$  by the map

$$\vartheta \rightarrow \varphi_T^{-1}(\vartheta - \vartheta_0)$$

and by  $\hat{\mu}'_T$  the maximum likelihood estimator

$$\hat{\mu}'_T = \arg \max_{\mu \in \mathfrak{M}_T} \ell_T(\mu).$$

---

<sup>3)</sup> This remark and the subsequent ones have to be modified in an obvious way in the periodic Case P1.



Then

$$\hat{\vartheta}'_T = \vartheta_0 + \varphi_T \hat{\mu}'_T,$$

Case 1: It is easy to see that  $\mathfrak{M}_T$  is a straight line through  $(0, 0)$  with a slope tending to zero if  $T \rightarrow \infty$ . So we obtain with the notation  $\mu = (\alpha, \beta)^*$

$$\hat{\mu}'_T \xrightarrow{d} \arg \max_{\beta=0, \alpha \in R_1} (\mu^* V_\infty - \frac{1}{2} \mu^* I_\infty \mu). \quad (3.5)$$

This means that the limit distribution of  $\hat{\mu}'_T$  is the distribution of the vector  $\Phi(1, 0)^*$ , where  $\Phi = V_{\infty,1}/I_{\infty,11}$  ( $V_{\infty,i}$  are the elements of  $V_\infty$  and  $I_{\infty,ii}$  are the diagonal elements of  $I_\infty$ ,  $i = 1, 2$ ). Since  $\varphi_{11}^{-1}(T)(1, e^{-v_0})^* \varphi_T \rightarrow (1, 0)^*$ , we get

$$\varphi_{11}^{-1}(T)(1, e^{-v_0})^*(\hat{\mu}'_T - \vartheta_0) \xrightarrow{d} \Phi,$$

hence

$$\varphi_{11}^{-1}(T)(\hat{\vartheta}'_T - \vartheta_0) \xrightarrow{d} \Phi(1 + c, -e^{v_0}c)^*$$

for some real  $c$ . Note that here the rate of convergence of  $\hat{\vartheta}'_T$  to  $\vartheta_0$  equals  $\varphi_{11}^{-1}(T)$ .

Case 2: Let us assume additionally that we are not in Case M3. Then  $\mathfrak{M}_T = \{(\alpha, \beta) : \alpha = 0\}$  and we obtain

$$\hat{\mu}'_T \xrightarrow{d} \arg \max_{\alpha=0, \beta \in R_1} (\mu^* V_\infty - \frac{1}{2} \mu^* I_\infty \mu).$$

Thus the limit distribution of  $\hat{\mu}'_T$  is the distribution of the vector  $\Psi(0, 1)^*$ , where  $\Psi = V_{\infty,2}/I_{\infty,22}$ . Now it is easy to see that

$$\varphi_{22}^{-1}(T)(\hat{\mu}'_T - \vartheta_0) \xrightarrow{d} \Psi(1, -e^{v_0})^* \quad (3.6)$$

Therefore, the rate of convergence of  $\hat{\vartheta}'_T$  to  $\vartheta_0$  is  $\varphi_{22}^{-1}(T)$ . Moreover, if  $I_\infty$  is diagonal (this happens in Cases M1, Q1, Q4, and Q5) then

$$\varphi_{22}^{-1}(T)(\hat{\vartheta}'_T - \hat{\vartheta}'_T) \xrightarrow{P} 0 \quad (3.7)$$

(compare (3.4) and (3.6)). Furthermore,  $I_{\infty,22}$  is nonrandom in Cases M1 and Q1. Hence the submodel  $(P_\vartheta, \vartheta \in \Theta_0)$  has the LAN property in these two cases. Applying the asymptotic minimax theorem to this submodel and taking into account (3.7), we obtain the following local asymptotic minimax bound for an arbitrary estimator  $\hat{\vartheta}_T$ :

$$\lim_{r \rightarrow \infty} \liminf_{T \rightarrow \infty} \sup_{T^{1/2} \|\hat{\vartheta}_T - \vartheta_0\| \leq r} E_\vartheta w\{T^{1/2} \|\hat{\vartheta}_T - \vartheta\|\} \geq E w\{|\sigma|^{-1} (1 + e^{2v_0})^{1/2} Z\},$$

where  $Z$  is a standard normal variable,  $\sigma^2 = \int_0^\infty (x_0(t) - e^{v_0} x_0(t-1))^2 dt$ ,  $w : R^1 \rightarrow [0, \infty)$  is a bowl shape loss function; here  $\vartheta_0$  satisfies  $a < 1$ ,  $-a \leq b < w(a)$  or  $a \geq 1$ ,  $-a < b < w(a)$  (this corresponds to Cases M1 and Q1). Note that a similar estimate can be obtained from (3.3) also in Case N. The maximum likelihood estimator  $\hat{\vartheta}_T$  attains this bound, at least for bounded  $w$ .

Finally, let us consider Case M3. Here we have  $\varphi_{11}(T) = T^{-1} e^{-v_0 T}$  and  $\varphi_{22}(T) = T e^{-v_0 T}$ . Thus  $\mathfrak{M}_T$  is the straight line through  $(0, 0)$  with the slope  $1/T$ .  $\hat{\mu}'_T$  has the limit distribution as in (3.5) above. This implies

$$e^{v_0 T}(\hat{\vartheta}'_T - \vartheta_0) \xrightarrow{d} \Phi(1, -e^{v_0})^*.$$

So here the rate of convergence of  $\hat{\vartheta}'_T$  to  $\vartheta_0$  is intermediate between  $\varphi_{11}^{-1}(T)$  and  $\varphi_{22}^{-1}(T)$ .

## 4 Proofs

The main goal of this section is to prove Propositions 2.1–2.11, i.e. to prove the weak convergence of  $(V_T, I_T)$  to the corresponding limit. Unless otherwise specified, all limits are taken as  $T \rightarrow \infty$ .

Let us start with some general remarks. With the exception of Case M3, we have

$$V_T = \left( \varphi_{11}(T) \int_0^T X(t) dW(t), \varphi_{22}(T) \int_0^T Y(t) dW(t) \right)^* \quad (4.1)$$

and

$$I_T = \begin{pmatrix} \varphi_{11}^2(T) \int_0^T X^2(t) dt & \varphi_{11}(T) \varphi_{22}(T) \int_0^T X(t) Y(t) dt \\ \varphi_{11}(T) \varphi_{22}(T) \int_0^T X(t) Y(t) dt & \varphi_{22}^2(T) \int_0^T Y^2(t) dt \end{pmatrix}, \quad (4.2)$$

where

$$Y(t) = \begin{cases} X(t-1) & \text{in Cases N, P2, Q3,} \\ X(t) - e^{v_0} X(t-1) & \text{otherwise.} \end{cases} \quad (4.3)$$

Note that the process  $Y(t)$  defined in this way has a representation similar to (1.2), where the function  $x_0(t)$  is replaced by the corresponding to (4.3) linear combination of  $x_0(t)$  and  $x_0(t-1)$ ; this representation holds for  $t \geq 1$ .

More generally, let us say that a continuous process  $(Y(t), t \geq 0)$  has the representation (4.4) below with a function  $y(\cdot)$  if

$$Y(t) = y(t)X_0(0) + b \int_{-1}^0 y(t-s-1)X_0(s)ds + \int_0^t y(t-s)dW(s), \quad t \geq t_0, \quad (4.4)$$

where  $t_0 \geq 1$  and  $y = (y(t), t \geq 0)$  is a deterministic continuous function. Before proving Propositions 2.1–2.11, we shall study some properties of processes having the representation (4.4).

Our first lemma summarizes in an appropriate form some simple facts being used over and over throughout this section. The proof is trivial and therefore omitted.

**Lemma 4.1** *Assume  $Y_1(t)$ ,  $Y_2(t)$  and  $Z(t)$ ,  $t \geq 0$ , are adapted continuous processes,  $Y(t) = Y_1(t) + Y_2(t)$ ,  $t \geq 0$ , and  $(W(t), t \geq 0)$  is a standard Wiener process. Moreover let  $\delta(T)$  and  $\varepsilon(T)$  be normalizing functions such that*

$$\delta^2(T) \int_0^T Y_1^2(t) dt, \quad T \geq 0, \quad \text{and} \quad \varepsilon^2(T) \int_0^T Z^2(t) dt, \quad T \geq 0,$$

*are bounded in probability and*

$$\delta^2(T) \int_0^T Y_2^2(t) dt \xrightarrow{P} 0.$$

Then

$$\begin{aligned}\delta(T) \left\{ \int_0^T Y(t) dW(t) - \int_0^T Y_1(t) dW(t) \right\} &\xrightarrow{P} 0, \\ \delta^2(T) \left\{ \int_0^T Y^2(t) dt - \int_0^T Y_1^2(t) dt \right\} &\xrightarrow{P} 0. \\ \delta(T)\varepsilon(T) \left\{ \int_0^T Y(t)Z(t) dt - \int_0^T Y_1(t)Z(t) dt \right\} &\xrightarrow{P} 0.\end{aligned}$$

Let  $(Y(t), t \geq 0)$  be a process having the representation (4.4). Sometimes the first term in the right-hand side of (4.4) is small in the sense of Lemma 4.1, i.e. it can be chosen as  $Y_2(t)$ . The next lemma shows that then the second term in the right-hand side of (4.4) is also small in the same sense.

**Lemma 4.2** *Put*

$$z(t) = \int_{-1}^0 y(t-s-1)X_0(s) ds, \quad t \geq t_0.$$

Then

$$\int_{t_0}^T z^2(t) dt \leq \int_{-1}^0 X_0^2(s) ds \int_0^T y^2(t) dt.$$

For the proof use Fubini's theorem and the Cauchy-Schwartz inequality.

In Lemmas 4.3, 4.5, 4.7, 4.8 and Corollary 4.4 we assume that  $Y(\cdot)$ ,  $Y_1(\cdot)$ ,  $Y_2(\cdot)$  are continuous processes having the representation (4.4) with functions  $y(\cdot)$ ,  $y_1(\cdot)$ ,  $y_2(\cdot)$  respectively.

**Lemma 4.3** *Assume that  $y = (y(t), t \geq 0)$  is a square integrable function. Then*

$$\begin{aligned}T^{-1} \int_0^T Y(t) dt &\xrightarrow{P} 0, \\ T^{-1} \int_0^T Y^2(t) dt &\xrightarrow{P} \sigma^2 := \int_0^\infty y^2(t) dt.\end{aligned}$$

**Proof:** According to Lemmas 4.1 and 4.2, it is sufficient to prove the assertion for  $X_0(s) \equiv 0$ . We introduce the stationary process  $Z(t) = \int_{-\infty}^t y(t-s) dW(s)$ ,  $t \geq 0$ , where  $W(\cdot)$  is independently of  $(W(s), s \geq 0)$  extended to  $(-\infty, 0)$  as a Wiener process.

Obviously, we have

$$T^{-1} E \int_{t_0}^T (Z(t) - Y(t))^2 dt = T^{-1} \int_{t_0}^T \int_t^\infty y^2(s) ds dt \rightarrow 0. \quad (4.5)$$

Applying the law of large numbers to the Gaussian stationary process  $Z(\cdot)$ , which is ergodic, we get

$$T^{-1} \int_0^T Z(t) dt \xrightarrow{P} EZ(0) = 0, \quad T^{-1} \int_0^T Z^2(t) dt \xrightarrow{P} EZ^2(0) = \sigma^2.$$

Now the claim follows from (4.5).

**Corollary 4.4** *If  $\int_0^\infty y_i^2(t) dt < \infty$ ,  $i = 1, 2$ , then*

$$T^{-1} \int_0^T Y_1(t) Y_2(t) dt \xrightarrow{P} \int_0^\infty y_1(t) y_2(t) dt.$$

**Lemma 4.5** *Suppose that  $y(t) = t^\alpha e^{wt}$  for some  $\alpha = 0, 1, 2, \dots$  and  $w > 0$ . Then with probability one*

$$\lim_{t \rightarrow \infty} t^{-\alpha} e^{-wt} Y(t) = U,$$

where

$$U = X_0(0) + b \int_{-1}^0 e^{-w(s+1)} X_0(s) ds + \int_0^\infty e^{-ws} dW(s).$$

**Proof:** Using the representation (4.4) of  $Y(\cdot)$  the appearance of the first two terms of  $U$  is quite obvious. Furthermore,

$$\begin{aligned} t^{-\alpha} e^{-wt} \int_0^t y(t-s) dW(s) &= \int_0^t \left(1 - \frac{s}{t}\right)^\alpha e^{-ws} dW(s) \\ &= \int_0^t e^{-ws} dW(s) \\ &\quad + \sum_{k=1}^{\alpha} (-1)^k \binom{\alpha}{k} t^{-k} \int_0^t s^k e^{-ws} dW(s). \end{aligned}$$

It remains to note that with probability one

$$\lim_{t \rightarrow \infty} \int_0^t e^{-ws} dW(s) = \int_0^\infty e^{-ws} dW(s)$$

by Lévy's theorem and

$$\lim_{t \rightarrow \infty} t^{-k} \int_0^t s^k e^{-ws} dW(s) = 0$$

by the strong law of large numbers for martingales, see e.g. Liptser and Shiryaev (1989, Chapter 2, § 6, Theorem 10).

**Lemma 4.6** *Let  $Z_1(\cdot)$  and  $Z_2(\cdot)$  be two continuous processes such that with probability one*

$$\lim_{t \rightarrow \infty} t^{-\alpha_i} e^{-w_i t} Z_i(t) = U_i,$$

for some  $\alpha_i \in \mathbb{R}$ ,  $w_i > 0$ , and some random variables  $U_i$  being finite almost surely,  $i = 1, 2$ . Then

$$T^{-\alpha_1 - \alpha_2} e^{-(w_1 + w_2)T} \int_0^T Z_1(t) Z_2(t) dt \xrightarrow{P} \frac{U_1 U_2}{w_1 + w_2}$$

and

$$T^{-\alpha_1 - 1/2} e^{-w_1 T} \int_0^T |Z_1(t)| dt \xrightarrow{P} 0.$$

**Remark:** In fact, we have the almost sure convergence in the assertions of Lemma. For the proof apply the L'Hospital rule.

**Lemma 4.7** Suppose that a continuous process  $Y(\cdot)$  has the representation (4.4) with a bounded  $y(\cdot)$ . If  $Z(\cdot)$  is a continuous process such that with probability one

$$\lim_{t \rightarrow \infty} e^{-wt} Z(t) = U$$

for some  $w > 0$  and some random variable  $U$  being finite almost surely, then

$$T^{-1} e^{-wT} \int_0^T Y(t) Z(t) dt \xrightarrow{P} 0.$$

If, moreover,  $y(\cdot)$  is square integrable on  $[0, \infty)$  then this convergence holds for  $T^{-1/2}$  instead of  $T^{-1}$ .

**Proof:** According to Lemmas 4.1 and 4.2 we can assume that  $X_0(0) \equiv 0$ . Applying these lemmas again we can substitute  $Z(t)$  by  $e^{wt}U$ . Thus it remains to prove that

$$T^{-1} e^{-wT} \int_0^T e^{wt} Y(t) dt \xrightarrow{P} 0.$$

(or  $T^{-1/2}$  instead of  $T^{-1}$  if  $\int_0^\infty y^2(t) dt < \infty$ ).

Now observe that

$$\begin{aligned} E \left| \int_0^T e^{wt} Y(t) dt \right| &\leq \int_0^T e^{wt} E |Y(t)| dt \\ &\leq \int_0^T e^{wt} (E |Y(t)|^2)^{1/2} dt \end{aligned}$$

and that

$$E |Y(t)|^2 = \int_0^t y^2(s) ds, \quad t \geq t_0,$$

which implies the assertion.

**Lemma 4.8** Assume that  $y(t) = \phi(t)e^{wt}$ , where  $\phi(t) = \cos(\xi t)$  or  $\phi(t) = \sin(\xi t)$  and  $w > 0$ . Then with probability one

$$\lim_{t \rightarrow \infty} \{e^{-wt} Y(t) - U(t)\} = 0,$$

where

$$U(t) = X_0(0)\phi(t) + b \int_{-1}^0 \phi(t-s-1) e^{-w(s+1)} X_0(s) ds + \int_0^\infty \phi(t-s) e^{-ws} dW(s)$$

is a continuous periodic process.

**Proof:** Note that

$$e^{-wt} Y(t) - U(t) = - \int_t^\infty \phi(t-s) e^{-ws} dW(s), \quad t \geq t_0.$$

If  $\phi(t) = \cos(\xi t)$  then

$$\int_t^\infty \phi(t-s) e^{-ws} dW(s) = \cos(\xi t) \int_t^\infty \cos(\xi s) e^{-ws} dW(s) + \sin(\xi t) \int_t^\infty \sin(\xi s) e^{-ws} dW(s),$$

which obviously tends almost surely to zero. The case  $\phi(t) = \sin(\xi t)$  can be treated similarly.

**Lemma 4.9** *Let  $Z(\cdot)$  be a continuous process such that with probability one*

$$\lim_{t \rightarrow \infty} \{e^{-wt} Z(t) - U(t)\} = 0,$$

*where  $U(\cdot)$  is a continuous periodic process on  $R$  and  $w > 0$ . Then*

$$e^{-2wT} \int_0^T Z^2(t) dt - \int_0^\infty e^{-2wt} U^2(T-t) dt \xrightarrow{P} 0.$$

**Proof:** Applying Lemma 4.1 we can replace  $Z(t)$  by  $e^{wt}U(t)$  and observe that

$$e^{-2wT} \int_0^T e^{2wt} U^2(t) dt = \int_0^T e^{-2wt} U^2(T-t) dt$$

and

$$\int_T^\infty e^{-2wt} U^2(T-t) dt \xrightarrow{P} 0.$$

**Lemma 4.10** *Let  $Z_1(\cdot)$  and  $Z_2(\cdot)$  be two continuous processes such that with probability one*

$$\lim_{t \rightarrow \infty} e^{-w_1 t} Z_1(t) = U_1$$

*and*

$$\lim_{t \rightarrow \infty} \{e^{-w_2 t} Z_2(t) - U_2(t)\} = 0,$$

*where  $w_1, w_2 > 0$ ,  $U_1$  is a finite random variable,  $U_2(\cdot)$  is a continuous periodic process on  $R$ . Then*

$$e^{-(w_1+w_2)T} \int_0^T Z_1(t) Z_2(t) dt - U_1 \int_0^\infty e^{-(w_1+w_2)t} U_2(T-t) dt \xrightarrow{P} 0.$$

**Proof:** The proof is analogous to that of Lemma 4.9.

Now we are in a position to prove Propositions 2.1–2.8.

**Proof of Proposition 2.1:** According to (4.1)–(4.3),

$$V_T = \left( T^{-1/2} \int_0^T X(t) dW(t), T^{-1/2} \int_0^T X(t-1) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} T^{-1} \int_0^T X^2(t) dt & T^{-1} \int_0^T X(t) X(t-1) dt \\ T^{-1} \int_0^T X(t) X(t-1) dt & T^{-1} \int_0^T X^2(t-1) dt \end{pmatrix}.$$

The process  $(X(t), t \geq 0)$  has the representation (4.4) with  $y(t) = x_0(t)$ ,  $t \geq 0$ , and  $(X(t-1), t \geq 0)$  has this representation with  $y(t) = x_0(t-1)$ ,  $t \geq 0$ .

By assumption,  $v_0 < 0$ , i.e.  $\int_0^\infty x_0^2(t) dt < \infty$  holds. Thus we can apply Lemma 4.3 and Corollary 4.4 to obtain  $I_T \xrightarrow{P} I_\infty$ . Now the claim follows from the central limit theorem for martingales.

**Proof of Proposition 2.2:** According to (4.1)–(4.3),

$$V_T = \left( e^{-v_0 T} \int_0^T X(t) dW(t), T^{-1/2} \int_0^T Y(t) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} e^{-2v_0 T} \int_0^T X^2(t) dt & T^{-1/2} e^{-v_0 T} \int_0^T X(t) Y(t) dt \\ T^{-1/2} e^{-v_0 T} \int_0^T X(t) Y(t) dt & T^{-1} \int_0^T Y^2(t) dt \end{pmatrix},$$

where

$$Y(t) = X(t) - e^{v_0} X(t-1), \quad t \geq 0. \quad (4.6)$$

Note that  $Y(\cdot)$  has the representation (4.4) with  $y(t) = x_0(t) - e^{v_0} x_0(t-1)$ . In the considered case  $v_0 > 0$  and  $v_1 < 0$ . It follows from Lemma 1.1 that

$$x_0(t) = \frac{1}{v_0 - a + 1} e^{v_0 t} + o(e^{\gamma t}) \quad (4.7)$$

for some  $\gamma < 0$ , and this implies  $y(t) = o(e^{\gamma t})$  and therefore  $\int_0^\infty y^2(t) dt < \infty$ . Now Lemmas 4.5, 4.6 and 4.1 imply

$$e^{-2v_0 T} \int_0^T X^2(t) dt \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2},$$

it follows from Lemmas 4.5, 4.7 and 4.1 that

$$T^{-1/2} e^{-v_0 T} \int_0^T X(t) Y(t) dt \xrightarrow{P} 0,$$

and Lemma 4.3 implies

$$T^{-1} \int_0^T Y^2(t) dt \xrightarrow{P} \int_0^\infty y^2(t) dt.$$

Summarizing these results we get the convergence in probability of  $I_T$  to  $I_\infty$ . The joint convergence of  $(V_T, I_T)$  to  $(V_\infty, I_\infty)$  follows from the stable limit theorem for martingales, see Jacod and Shiryaev (1987, Theorem VIII.5.42) or Touati (1991, Theorem 1).

**Proof of Proposition 2.3:** According to (4.1)–(4.3),

$$V_T = \left( e^{-v_0 T} \int_0^T X(t) dW(t), e^{-v_1 T} \int_0^T Y(t) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} e^{-2v_0 T} \int_0^T X^2(t) dt & e^{-(v_0+v_1)T} \int_0^T X(t) Y(t) dt \\ e^{-(v_0+v_1)T} \int_0^T X(t) Y(t) dt & e^{-2v_1 T} \int_0^T Y^2(t) dt \end{pmatrix},$$

where  $Y(t)$  is defined as in (4.6) above and has the representation (4.4) with  $y(t) = x_0(t) - e^{v_0}x_0(t-1)$ . As in the previous proposition, it is sufficient to check that  $I_T \xrightarrow{P} I_\infty$ .

Since  $v_0 > v_1 > 0$ ,  $v_0 \in \Lambda$  and  $v_1 \in \Lambda$  in the considered case, it follows from Lemma 1.1 and (1.12) that (4.7) holds for some  $\gamma < v_0$  and

$$y(t) = \frac{e^{v_0-v_1} - 1}{a - v_1 - 1} e^{v_1 t} + o(e^{\gamma t})$$

for some  $\gamma_1 < v_1$ .

Using Lemmas 4.5, 4.6 and 4.1, we get

$$\begin{aligned} e^{-2v_0 T} \int_0^T X^2(t) dt &\xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2}, \\ e^{-(v_0+v_1)T} \int_0^T X(t)Y(t) dt &\xrightarrow{P} \frac{U_0 U_1 (e^{v_0-v_1} - 1)}{(v_0 + v_1)(v_0 - a + 1)(a - v_1 - 1)} \end{aligned}$$

and

$$e^{-2v_1 T} \int_0^T Y^2(t) dt \xrightarrow{P} \frac{U_1^2 (e^{v_0-v_1} - 1)^2}{2v_1(a - v_1 - 1)^2},$$

which yields the desired convergence.

**Proof of Proposition 2.4:** By the choice of  $\varphi_T$ , we have

$$V_T = \left( T^{-1} e^{-v_0 T} \int_0^T X(t) dW(t), T e^{-v_0 T} \int_0^T (Y(t) - T^{-1} Z(t)) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} T^{-2} e^{-2v_0 T} \int_0^T X^2(t) dt & e^{-2v_0 T} \int_0^T X(t)(Y(t) - T^{-1} Z(t)) dt \\ e^{-2v_0 T} \int_0^T X(t)(Y(t) - T^{-1} Z(t)) dt & T^2 e^{-2v_0 T} \int_0^T (Y(t) - T^{-1} Z(t))^2 dt \end{pmatrix},$$

where  $Z(t) = e^{v_0} X(t-1)$  and  $Y(t)$  is defined as in (4.6) above.

Obviously  $Z(\cdot)$  has the representation (4.4) with  $z(t) = e^{v_0} x(t-1)$ .

It follows from (1.11) that

$$x_0(t) = (2t + \frac{2}{3})e^{v_0 t} + o(e^{\gamma t}),$$

$$y(t) = 2e^{v_0 t} + o(e^{\gamma t}),$$

$$z(t) = (2t - \frac{4}{3})e^{v_0 t} + o(e^{\gamma t})$$

for some  $0 < \gamma < v_0$ .

Put

$$\hat{y}(t) = 2e^{v_0 t}, \quad \hat{z}(t) = (2t - \frac{4}{3})e^{v_0 t},$$



and let  $\widehat{Y}(t)$  and  $\widehat{Z}(t)$  be continuous processes having the representation (4.4) with the functions  $\widehat{y}(t)$  and  $\widehat{z}(t)$  respectively,  $\widehat{X}(t) = \widehat{Y}(t) + \widehat{Z}(t)$ . It can be easily checked that

$$e^{-2\gamma T} \int_0^T (X(t) - \widehat{X}(t))^2 dt \xrightarrow{P} 0 \quad (4.8)$$

$$e^{-2\gamma T} \int_0^T (Y(t) - \widehat{Y}(t))^2 dt \xrightarrow{P} 0 \quad (4.9)$$

and

$$e^{-2\gamma T} \int_1^T t^{-2} (Z(t) - \widehat{Z}(t))^2 dt \xrightarrow{P} 0. \quad (4.10)$$

Lemma 4.5 implies that with probability one

$$\lim_{t \rightarrow \infty} t^{-1} e^{-v_0 t} \widehat{X}(t) = 2U_0 \quad (4.11)$$

and

$$\lim_{t \rightarrow \infty} t^{-1} e^{-v_0 t} \widehat{Z}(t) = 2U_0; \quad (4.12)$$

the same proof as in Lemma 4.5 shows that

$$\lim_{t \rightarrow \infty} t e^{-v_0 t} (\widehat{Y}(t) - t^{-1} \widehat{Z}(t)) = \frac{4}{3}U_0 + 2U_2. \quad (4.13)$$

By Lemma 4.6 we obtain

$$T^{-2} e^{-2v_0 T} \int_0^T \widehat{X}^2(t) dt \xrightarrow{P} \frac{2U_0^2}{v_0}, \quad (4.14)$$

$$e^{-2v_0 T} \int_1^T \widehat{X}(t) (\widehat{Y}(t) - t^{-1} \widehat{Z}(t)) dt \xrightarrow{P} \frac{U_0(\frac{4}{3}U_0 + 2U_2)}{v_0}, \quad (4.15)$$

and

$$T^2 e^{-2v_0 T} \int_1^T (\widehat{Y}(t) - t^{-1} \widehat{Z}(t))^2 dt \xrightarrow{P} \frac{(\frac{4}{3}U_0 + 2U_2)^2}{2v_0}. \quad (4.16)$$

It follows from (4.11)–(4.13) by the L'Hospital rule that

$$e^{-2v_0 T} \int_1^T \widehat{X}(t) \widehat{Z}(t) (t^{-1} - T^{-1}) dt \xrightarrow{P} \frac{U_0^2}{v_0^2}, \quad (4.17)$$

$$T^2 e^{-2v_0 T} \int_1^T \widehat{Z}^2(t) (t^{-1} - T^{-1})^2 dt \xrightarrow{P} \frac{U_0^2}{v_0^3}, \quad (4.18)$$

$$T^2 e^{-2v_0 T} \int_1^T (\widehat{Y}(t) - t^{-1} \widehat{Z}(t)) \widehat{Z}(t) (t^{-1} - T^{-1}) dt \xrightarrow{P} \frac{U_0(\frac{4}{3}U_0 + 2U_2)}{2v_0^2}. \quad (4.19)$$

It follows from (4.8)–(4.10) that we can replace  $\widehat{X}(t)$ ,  $\widehat{Y}(t)$ , and  $\widehat{Z}(t)$  by  $X(t)$ ,  $Y(t)$ , and  $Z(t)$  respectively in relations (4.14)–(4.19). This implies the convergence  $I_T \xrightarrow{P} I_\infty$ . Now the claim follows from the stable limit theorem for martingales as in Propositions 2.2 and 2.3.

**Proof of Proposition 2.5:** According to (4.1)–(4.3),

$$V_T = \left( e^{-v_0 T} \int_0^T X(t) dW(t), e^{-v_1 T} \int_0^T Y(t) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} e^{-2v_0 T} \int_0^T X^2(t) dt & e^{-(v_0+v_1)T} \int_0^T X(t)Y(t) dt \\ e^{-(v_0+v_1)T} \int_0^T X(t)Y(t) dt & e^{-2v_1 T} \int_0^T Y^2(t) dt \end{pmatrix},$$

where

$$Y(t) = X(t) - e^{v_0} X(t-1), \quad t \geq 0.$$

Note that  $Y(\cdot)$  has the representation (4.4) with  $y(t) = x_0(t) - e^{v_0} x_0(t-1)$ . It follows from (1.12) that

$$x_0(t) = \frac{1}{v_0 + 1 - a} e^{v_0 t} + \{A_1 \cos(\xi_1 t) + B_1 \sin(\xi_1 t)\} e^{v_1 t} + o(e^{\gamma t})$$

for some  $\gamma < v_1$  and, hence,

$$y(t) = \phi(t) e^{v_1 t} + o(e^{\gamma t}).$$

Applying Lemmas 4.5, 4.6, 4.8, 4.9, 4.10 and 4.1, we get

$$I_T - I_\infty(T) \xrightarrow{P} 0.$$

Now we complete the proof similarly to the previous case. The matrix-valued process  $I_\infty(T)$  is periodic with period  $\Delta = 2\pi/\xi_1$ , and the claim follows from the stable limit theorem for martingales as in previous propositions.

**Proof of Proposition 2.6:** According to (4.1)–(4.3),

$$V_T = \left( e^{-v_0 T} \int_0^T X(t) dW(t), e^{-v_0 T} \int_0^T X(t-1) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} e^{-2v_0 T} \int_0^T X^2(t) dt & e^{-2v_0 T} \int_0^T X(t)X(t-1) dt \\ e^{-2v_0 T} \int_0^T X(t)X(t-1) dt & e^{-2v_0 T} \int_0^T X^2(t-1) dt \end{pmatrix}.$$

The process  $(X(t), t \geq 0)$  has the representation (4.4) with  $y(t) = x_0(t)$ ,  $t \geq 0$ , and  $(X(t-1), t \geq 0)$  has this representation with  $y(t) = x_0(t-1)$ ,  $t \geq 0$ . It follows from (1.11) that

$$x_0(t) = \phi_0(t) e^{v_0 t} + o(e^{\gamma t})$$

for some  $\gamma < v_0$  because of  $\phi_0 = \psi_0$  by definition. Hence,

$$x_0(t-1) = \phi_2(t)e^{v_0 t} + o(e^{\gamma t}).$$

From Lemmas 4.8, 4.9 and 4.1 now we have

$$I_T - I_\infty(T) \xrightarrow{P} 0.$$

Obviously, the matrix-valued process  $I_\infty(T)$  is periodic with period  $\Delta = \pi/\xi_0$ , and we complete the proof similarly to the previous proposition.

**Proof of Proposition 2.7:** According to (4.1)–(4.3),

$$V_T = \left( T^{-1} \int_0^T X(t) dW(t), T^{-1/2} \int_0^T Y(t) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} T^{-2} \int_0^T X^2(t) dt & T^{-3/2} \int_0^T X(t)Y(t) dt \\ T^{-3/2} \int_0^T X(t)Y(t) dt & T^{-1} \int_0^T Y^2(t) dt \end{pmatrix},$$

where

$$Y(t) = X(t) - X(t-1), \quad t \geq 0.$$

Note that here we have

$$x_0(t) = \frac{1}{1-a} + o(e^{\gamma t})$$

for some  $\gamma < 0$  and  $X(\cdot)$  and  $Y(\cdot)$  have the representation (4.4) with the functions  $x_0(t)$  and  $y(t) = x_0(t) - x_0(t-1)$  respectively. Obviously it holds  $y(t) = o(e^{\gamma t})$ .

Consider the processes

$$W^T(s) = T^{-1/2}W(Ts),$$

$$X^T(s) = (1-a)^{-1}T^{-1} \int_0^{Ts} W(t) dW(t) = (1-a)^{-1} \int_0^s W^T(t) dW^T(t),$$

$$Y^T(s) = T^{-1/2} \int_0^{Ts} Y(t) dW(t) = \int_0^s Y(Tt) dW^T(t), \quad s \in [0, 1].$$

These processes are square integrable continuous martingales. Since

$$\int_0^s Y^2(Tt) dt = T^{-1} \int_0^{Ts} Y^2(t) dt \xrightarrow{P} \sigma^2 s,$$

$$\int_0^s Y(Tt) dt = T^{-1} \int_0^{Ts} Y(t) dt \xrightarrow{P} 0$$

by Lemma 4.3, the functional central limit theorem for martingales implies that

$$(W^T, Y^T) \xrightarrow{\mathcal{L}} (\widetilde{W}, \sigma \widetilde{W}_1),$$

where  $(\widetilde{W}_1(t), t \in [0, 1])$  is a standard Wiener process independent of  $\widetilde{W}(\cdot)$ . Since  $X^T(s) = ((W^T(s))^2 - s)/2(1-a)$  by Itô's formula, we also have

$$(X^T, Y^T) \xrightarrow{\mathcal{L}} (\widetilde{X}, \sigma \widetilde{W}_1), \quad (4.20)$$

where

$$\widetilde{X}(s) = (1-a)^{-1} \int_0^s \widetilde{W}(t) d\widetilde{W}(t).$$

Moreover, the convergence (4.20) implies the joint functional convergence of  $(X^T, Y^T)$  together with their quadratic (co-)variations, see Jacod and Shiryaev (1987, Theorem VI.6.1). In particular,  $(\widehat{V}_T, \widehat{I}_T) \xrightarrow{d} (V_\infty, I_\infty)$ , where

$$\widehat{V}_T = (X^T(1), Y^T(1))^* = \left( (1-a)^{-1} T^{-1} \int_0^T W(t) dW(t), T^{-1/2} \int_0^T Y(t) dW(t) \right)^*$$

and

$$\widehat{I}_T = \begin{pmatrix} (1-a)^{-2} T^{-2} \int_0^T W^2(t) dt & (1-a)^{-2} T^{-3/2} \int_0^T W(t) Y(t) dt \\ (1-a)^{-2} T^{-3/2} \int_0^T W(t) Y(t) dt & T^{-1} \int_0^T Y^2(t) dt \end{pmatrix}.$$

But, evidently,  $V_T - \widehat{V}_T \xrightarrow{P} 0$  and  $I_T - \widehat{I}_T \xrightarrow{P} 0$  by Lemmas 4.1 and 4.2.

**Proof of Proposition 2.8:** According to (4.1)–(4.3),

$$V_T^* = \left( T^{-2} \int_0^T X(t) dW(t), T^{-1} \int_0^T Y(t) dW(t) \right) \quad (4.21)$$

and

$$I_T = \begin{pmatrix} T^{-4} \int_0^T X^2(t) dt & T^{-3} \int_0^T X(t) Y(t) dt \\ T^{-3} \int_0^T X(t) Y(t) dt & T^{-2} \int_0^T Y^2(t) dt \end{pmatrix}, \quad (4.22)$$

where

$$Y(t) = X(t) - X(t-1), \quad t \geq 0.$$

Here we have

$$x_0(t) = (2t + \frac{2}{3}) + o(e^{\gamma t})$$

for some  $\gamma < 0$  and  $X(\cdot)$  and  $Y(\cdot)$  have the representation (4.4) with the functions  $x_0(t)$  and  $y(t) = x_0(t) - x_0(t-1)$ . Obviously it holds  $y(t) = 2 + o(e^{\gamma t})$ .

Let  $\widehat{V}_T$  and  $\widehat{I}_T$  be defined by the expressions (4.21) and (4.22) respectively after replacing  $X(t)$  by  $\widehat{X}(t) = 2 \int_0^t (t-s) dW(s)$  and  $Y(t)$  by  $2W(t)$ . We have  $V_T - \widehat{V}_T \xrightarrow{P} 0$  and  $I_T - \widehat{I}_T \xrightarrow{P} 0$  by Lemmas 4.1 and 4.2. Now it remains to note that  $\widehat{X}(t) = 2 \int_0^t W(s) ds$  by Itô's formula and

$$(\widehat{V}_T, \widehat{I}_T) \stackrel{d}{=} (V_\infty, I_\infty) \quad \text{for all } T > 0$$

in view of the self-similarity of the Wiener process.

To prove the remaining propositions we need an additional result. In the next lemma and corollary for each integer  $n$  we consider a  $d$ -dimensional process  $M^n = (M_t^n)_{t \in [0,1]}$  on a stochastic basis  $(\Omega, \mathcal{F}, (\mathcal{F}_t^n)_{t \in [0,1]}, P)$ , whose components  $M^{n,i}$  are continuous local martingales;  $M_0^n = 0$ . We also consider a  $d$ -dimensional process  $M = (M_t)_{t \in [0,1]}$  with the same properties, on a stochastic basis  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,1]}, P)$ . We denote by  $N^n$  the  $R^d \otimes R^d$ -valued process whose components  $N^{n,ij}$  are defined as stochastic integrals  $N_t^{n,ij} = \int_0^t M_s^{n,i} dM_s^{n,j}$ ,  $t \in [0,1]$ , and we associate the process  $N$  with  $M$  similarly. For the notion of stable convergence we refer to Jacod and Shiryaev (1987, Chapter VIII, § 5c).

**Lemma 4.11** *Assume that*

- (i)  $M^n \xrightarrow{\mathcal{L}} M$ ;
- (ii) *for every finite subdivision  $\tau = \{0 = t_0 < t_1 < \dots < t_m = 1\}$  of  $[0,1]$ , the vectors  $(\xi^n, M_{t_1}^n, \dots, M_{t_m}^n)$  converge  $\mathcal{G}$ -stably to the vector  $(\xi, M_{t_1}, \dots, M_{t_m})$ , where  $\mathcal{G}$  is a sub- $\sigma$ -algebra of  $\mathcal{F}$ ,  $\xi^n$  and  $\xi$  are random variables.*

*Then, for every finite subdivision  $\tau$  of  $[0,1]$ , the vectors  $(\xi^n, M_{t_1}^n, \dots, M_{t_m}^n, N_{t_1}^n, \dots, N_{t_m}^n)$  converge  $\mathcal{G}$ -stably to the vector  $(\xi, M_{t_1}, \dots, M_{t_m}, N_{t_1}, \dots, N_{t_m})$ .*

**Proof:** The proof is an easy consequence of the following fact. Let  $\nu = \{0 = u_0 < u_1 < \dots < u_k = 1\}$  be a subdivision of  $[0,1]$ . Put

$$S_t^{n,ij}(\nu) = \sum_{p=0}^{k-1} M_{u_p}^{n,i} (M_{t \wedge u_{p+1}}^{n,j} - M_{t \wedge u_p}^{n,j}).$$

Then, for any  $\varepsilon > 0$ , there exists a  $\delta > 0$  such that for all subdivisions  $\nu$  of  $[0,1]$  satisfying  $|\nu| = \sup_{1 \leq p \leq k} (u_p - u_{p-1}) \leq \delta$ , we have

$$\sup_n P \left( \sup_{t \in [0,1]} |S_t^n(\nu) - N_t^n| \geq \varepsilon \right) \leq \varepsilon.$$

This can be shown from (i) following the lines of the proof of Lemma VI.6.13 in Jacod and Shiryaev (1987); moreover, the proof is much simpler in our case since  $M^{n,j}$  are assumed to be continuous local martingales.

**Corollary 4.12** *Let the assumptions of Lemma 4.11 be fulfilled. Denote by  $[N^n, N^n]$  the  $(R^d \otimes R^d) \otimes (R^d \otimes R^d)$ -valued process whose components are the quadratic covariations  $[N^{n,ij}, N^{n,kl}]$ ;  $[N, N]$  is defined similarly. Then the vectors  $(\xi^n, N_1^n, [N^n, N^n]_1)$  converge  $\mathcal{G}$ -stably to the vector  $(\xi, N_1, [N, N]_1)$ .*

**Proof:** Note that  $[N^{n,ij}, N^{n,kl}]_t = N_t^{n,ij} N_t^{n,kl} - \int_0^t N_s^{n,ij} dN_s^{n,kl} - \int_0^t N_s^{n,kl} dN_s^{n,ij}$  by Itô's formula, so the claim follows from Lemma 4.11 applied to the processes  $N^n$ .

**Remark:** If  $\xi^n \equiv \xi$  and  $\mathcal{G} = \{\emptyset, \Omega\}$ , the assertions of Lemma 4.11 and Corollary 4.12 are very special cases of theorems on convergence of stochastic integrals, see Jakubowski

*et al.* (1989) and Kurtz and Protter (1991), cf. also Jacod and Shiryaev (1987, Theorem VI.6.1).

**Proof of Proposition 2.9:** According to (4.1)–(4.3),

$$V_T = \left( T^{-1} \int_0^T X(t) dW(t), T^{-1} \int_0^T X(t-1) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} T^{-2} \int_0^T X^2(t) dt & T^{-2} \int_0^T X(t)X(t-1) dt \\ T^{-2} \int_0^T X(t)X(t-1) dt & T^{-2} \int_0^T X^2(t-1) dt \end{pmatrix}.$$

Because of (1.2) the process  $(X(t), t \geq 0)$  has the representation (4.4) with the function  $x_0(t)$ ,  $t \geq 0$ , and  $(X(t-1), t \geq 0)$  has this representation with the function  $y(t) = x_0(t-1)$ ,  $t \geq 0$ .

By Lemma 1.1 we have

$$x_0(t) = A_0 \cos(\xi_0 t) + B_0 \sin(\xi_0 t) + o(e^{\gamma t}) \quad (4.23)$$

and

$$x_0(t-1) = A_2 \cos(\xi_0 t) + B_2 \sin(\xi_0 t) + o(e^{\gamma t})$$

for some  $\gamma < 0$ . We introduce  $(X_1(t), t \geq 0)$  and  $(X_2(t), t \geq 0)$  by

$$X_1(t) = \int_0^t \cos(\xi_0 s) dW(s) \quad \text{and} \quad X_2(t) = \int_0^t \sin(\xi_0 s) dW(s).$$

The solution  $(X(t), t \geq 0)$  has the representation

$$X(t) = \int_0^t x_0(t-s) dW(s) + x_0(t)X_0(0) + b \int_{-1}^0 x_0(t-s-1)X_0(s) ds, \quad t \geq 0. \quad (4.24)$$

Inserting (4.23) into the first term we get

$$X(t) = A_0 \cos(\xi_0 t)X_1(t) + A_0 \sin(\xi_0 t)X_2(t) + B_0 \sin(\xi_0 t)X_1(t) - B_0 \cos(\xi_0 t)X_2(t) + \overline{X}(t),$$

where  $\overline{X}(t)$  is the sum of the last two terms in (4.24) and the contribution arising from the remainder term in (4.23);  $T^{-2} \int_0^T \overline{X}^2(t) dt \xrightarrow{P} 0$  by Lemmas 4.2 and 4.3.

Similarly, we get

$$X(t-1) = A_2 \cos(\xi_0 t)X_1(t) + A_2 \sin(\xi_0 t)X_2(t) + B_2 \sin(\xi_0 t)X_1(t) - B_2 \cos(\xi_0 t)X_2(t) + \overline{Y}(t),$$

where  $T^{-2} \int_0^T \overline{Y}^2(t) dt \xrightarrow{P} 0$ .

Consider the following processes on the interval  $[0, 1]$ :

$$W^T(s) = T^{-1/2}W(Ts),$$

$$X_1^T(s) = T^{-1/2}X_1(Ts) = \int_0^s \cos(\xi_0 Tt) dW^T(t),$$

$$X_2^T(s) = T^{-1/2} X_2(Ts) = \int_0^s \sin(\xi_0 T t) dW^T(t),$$

$$\begin{aligned} X^T(s) &= A_0 \cos(\xi_0 T s) X_1^T(s) + A_0 \sin(\xi_0 T s) X_2^T(s) \\ &\quad + B_0 \sin(\xi_0 T s) X_1^T(s) - B_0 \cos(\xi_0 T s) X_2^T(s), \end{aligned}$$

$$\begin{aligned} Y^T(s) &= A_2 \cos(\xi_0 T s) X_1^T(s) + A_2 \sin(\xi_0 T s) X_2^T(s) \\ &\quad + B_2 \sin(\xi_0 T s) X_1^T(s) - B_2 \cos(\xi_0 T s) X_2^T(s). \end{aligned}$$

Then

$$X(t) = T^{1/2} X^T(t/T) + \overline{X}(t), \quad X(t-1) = T^{1/2} Y^T(t/T) + \overline{Y}(t),$$

and by Lemma 4.1 it is enough to check that

$$(\widehat{V}_T, \widehat{I}_T) \xrightarrow{d} (V_\infty, I_\infty), \quad (4.25)$$

where

$$\widehat{V}_T = \left( \int_0^1 X^T(t) dW^T(t), \int_0^1 Y^T(t) dW^T(t) \right)^*$$

and

$$\widehat{I}_T = \begin{pmatrix} \int_0^1 \{X^T(t)\}^2 dt & \int_0^1 X^T(t) Y^T(t) dt \\ \int_0^1 X^T(t) Y^T(t) dt & \int_0^1 \{Y^T(t)\}^2 dt \end{pmatrix}.$$

But

$$\int_0^1 X^T dW^T = A_0 \int_0^1 X_1^T dX_1^T + A_0 \int_0^1 X_2^T dX_2^T + B_0 \int_0^1 X_1^T dX_2^T - B_0 \int_0^1 X_2^T dX_1^T$$

and

$$\int_0^1 Y^T dW^T = A_2 \int_0^1 X_1^T dX_1^T + A_2 \int_0^1 X_2^T dX_2^T + B_2 \int_0^1 X_1^T dX_2^T - B_2 \int_0^1 X_2^T dX_1^T$$

are represented as linear combinations of the stochastic integrals  $\int X_i^T dX_j^T$ ,  $i, j = 1, 2$ . Since

$$(X_1^T, X_2^T) \xrightarrow{\mathcal{L}} \frac{1}{\sqrt{2}}(\widetilde{W}_1, \widetilde{W}_2)$$

by the functional central limit theorem, the claim follows from Corollary 4.12.

**Proof of Proposition 2.10:** According to (4.1)–(4.3),

$$V_T = \left( e^{-v_0 T} \int_0^T X(t) dW(t), T^{-1} \int_0^T Y(t) dW(t) \right)^*$$

and

$$I_T = \begin{pmatrix} e^{-2v_0T} \int_0^T X^2(t) dt & T^{-1}e^{-v_0T} \int_0^T X(t)Y(t) dt \\ T^{-1}e^{-v_0T} \int_0^T X(t)Y(t) dt & T^{-2} \int_0^T Y^2(t) dt \end{pmatrix},$$

where

$$Y(t) = X(t) - e^{v_0}X(t-1), \quad t \geq 0,$$

and  $Y(\cdot)$  has the representation (4.4) with  $y(t) = x_0(t) - e^{v_0}x_0(t-1)$ . It follows from (1.12) that

$$x_0(t) = \frac{1}{v_0 - a + 1} e^{v_0 t} - \frac{1}{a - 1} + o(e^{\gamma t})$$

and

$$y(t) = \frac{e^{v_0} - 1}{a - 1} + o(e^{\gamma t})$$

for some  $\gamma < 0$ .

Due to Lemmas 4.5, 4.6, 4.7 and 4.1, this implies

$$e^{-2v_0T} \int_0^T X^2(t) dt \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2}, \quad (4.26)$$

$$T^{-1/2}e^{-v_0T} \int_0^T |X(t)| dt \xrightarrow{P} 0, \quad (4.27)$$

and

$$T^{-1}e^{-v_0T} \int_0^T X(t)Y(t) dt \xrightarrow{P} 0. \quad (4.28)$$

Introduce the following processes on the interval  $[0, 1]$ :

$$W^T(s) = T^{-1/2}W(Ts),$$

$$X^T(s) = e^{-v_0T} \int_0^{Ts} X(t) dW(t) = T^{1/2}e^{-v_0T} \int_0^s X(Tt) dW^T(t),$$

$$Y^T(s) = T^{-1} \int_0^{Ts} W(t) dW(t) = \int_0^s W^T(t) dW^T(t),$$

which are continuous local martingales with respect to the filtration  $\mathcal{F}_s^T = \sigma\{X_0(t), t \in [-1, 0]; W(t), t \in [0, Ts]\}$ .

Let  $\tau = \{0 = t_0 < t_1 < \dots < t_m = 1\}$  be a subdivision of  $[0, 1]$ . It follows from (4.26) and (4.27) that

$$T e^{-2v_0T} \int_0^1 X^2(Tt) dt \xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2}$$

and

$$T^{1/2}e^{-v_0T} \int_0^s X(Tt) dt \xrightarrow{P} 0, \quad s \in [0, 1],$$

therefore, we can apply the stable limit theorem for martingales (Jacod and Shiryaev 1987, Theorem VIII.5.42; or Touati 1991, Theorem 1) to the process  $X^T(\cdot)$  and to the stopped processes  $W^T(t_1 \wedge \cdot), \dots, W^T(t_m \wedge \cdot)$ , which yields that the vectors



$(X^T(1), W^T(t_1), \dots, W^T(t_m))$  converge  $\mathcal{F}$ -stably (where  $\mathcal{F} = \sigma\{X_0(t), t \in [-1, 0]; W(t), t \geq 0\}$ ) to the vector

$$\left( \frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \widetilde{W}(t_1), \dots, \widetilde{W}(t_m) \right)$$

as  $T \rightarrow \infty$ . Clearly,  $W^T \xrightarrow{\mathcal{L}} \widetilde{W}$ . Applying Corollary 4.12, we obtain the  $\mathcal{F}$ -stable convergence of the vector

$$\left( e^{-v_0 T} \int_0^T X(t) dW(t), T^{-1} \int_0^T W(t) dW(t), T^{-2} \int_0^T W^2(t) dt \right)$$

to the vector

$$\left( \frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \int_0^1 \widetilde{W}(t) d\widetilde{W}(t), \int_0^1 \widetilde{W}^2(t) dt \right).$$

By Lemmas 4.1 and 4.2, we have the  $\mathcal{F}$ -stable convergence of the vector

$$\left( e^{-v_0 T} \int_0^T X(t) dW(t), T^{-1} \int_0^T Y(t) dW(t), T^{-2} \int_0^T Y^2(t) dt \right)$$

to the vector

$$\left( \frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{e^{v_0} - 1}{a - 1} \int_0^1 \widetilde{W}(t) d\widetilde{W}(t), \frac{(e^{v_0} - 1)^2}{(a - 1)^2} \int_0^1 \widetilde{W}^2(t) dt \right).$$

Now the convergence  $(V_T, I_T) \xrightarrow{d} (V_\infty, I_\infty)$  follows from the properties of the stable convergence and relations (4.26) and (4.28).

**Proof of Proposition 2.11:** The proof follows the same lines as the proof of Proposition 2.10. Here we have

$$x_0(t) = \frac{1}{v_0 - a + 1} e^{v_0 t} + A_1 \cos(\xi_1 t) + B_1 \sin(\xi_1 t) + o(e^{\gamma t})$$

and

$$y(t) = A \cos(\xi_1 t) + B \sin(\xi_1 t) + o(e^{\gamma t})$$

for some  $\gamma < 0$ , in particular, (4.26), (4.27) and (4.28) are still true.

Introduce the processes  $W^T(s)$  and  $X^T(s)$ ,  $s \in [0, 1]$ , as in the proof of Proposition 2.10 and the processes

$$X_1^T(s) = \int_0^s \cos(\xi_1 T t) dW^T(t), \quad X_2^T(s) = \int_0^s \sin(\xi_1 T t) dW^T(t),$$

$$Y^T(s) = A \cos(\xi_1 T s) X_1^T(s) + A \sin(\xi_1 T s) X_2^T(s) + B \sin(\xi_1 T s) X_1^T(s) - B \cos(\xi_1 T s) X_2^T(s).$$

Note that

$$\int_0^1 Y^T dW^T = A \int_0^1 X_1^T dX_1^T + A \int_0^1 X_2^T dX_2^T + B \int_0^1 X_1^T dX_2^T - B \int_0^1 X_2^T dX_1^T \quad (4.29)$$

and

$$(X_1^T, X_2^T) \xrightarrow{\mathcal{L}} \frac{1}{\sqrt{2}}(\widetilde{W}_1, \widetilde{W}_2). \quad (4.30)$$

In view of (4.26) and (4.27), we have

$$\begin{aligned} T e^{-2v_0 T} \int_0^1 X^2(Tt) dt &\xrightarrow{P} \frac{U_0^2}{2v_0(v_0 - a + 1)^2}, \\ T^{1/2} e^{-v_0 T} \int_0^s X(Tt) \cos(\xi_1 Tt) dt &\xrightarrow{P} 0, \\ T^{1/2} e^{-v_0 T} \int_0^s X(Tt) \sin(\xi_1 Tt) dt &\xrightarrow{P} 0, \end{aligned}$$

where  $s \in [0, 1]$ . Let us again apply the stable limit theorem for martingales but now to the process  $X^T(\cdot)$  and to the stopped processes  $X_1^T(t_1 \wedge \cdot), \dots, X_1^T(t_m \wedge \cdot), X_2^T(t_1 \wedge \cdot), \dots, X_2^T(t_m \wedge \cdot)$ , where  $\tau = \{0 = t_0 < t_1 < \dots < t_m = 1\}$  is a subdivision of  $[0, 1]$ , which yields the  $\mathcal{F}$ -stable convergence of the vectors

$$(X^T(1), X_1^T(t_1), \dots, X_1^T(t_m), X_2^T(t_1), \dots, X_2^T(t_m))$$

to the vector

$$\left( \frac{U_0 Z}{\sqrt{2v_0(v_0 - a + 1)}}, \frac{1}{\sqrt{2}}\widetilde{W}_1(t_1), \dots, \frac{1}{\sqrt{2}}\widetilde{W}_1(t_m), \frac{1}{\sqrt{2}}\widetilde{W}_2(t_1), \dots, \frac{1}{\sqrt{2}}\widetilde{W}_2(t_m) \right).$$

In view of (4.29) and (4.30), applying Corollary 4.12, we obtain the  $\mathcal{F}$ -stable convergence of the vector

$$\left( X^T(1), \int_0^1 Y^T(t) dW^T(t), \int_0^1 \{Y^T(t)\}^2 dt \right)$$

to the vector

$$\begin{aligned} &\left( \frac{U_0 Z}{\sqrt{2v_0(v_0 - a + 1)}}, \frac{1}{2} \left( A \int_0^1 \widetilde{W}_1 d\widetilde{W}_1 + A \int_0^1 \widetilde{W}_2 d\widetilde{W}_2 + B \int_0^1 \widetilde{W}_1 d\widetilde{W}_2 - B \int_0^1 \widetilde{W}_2 d\widetilde{W}_1 \right), \right. \\ &\quad \left. \frac{1}{4}(A^2 + B^2) \int_0^1 (\widetilde{W}_1^2 + \widetilde{W}_2^2) dt \right). \end{aligned}$$

But  $X^T(1) = e^{-v_0 T} \int_0^T X(t) dW(t)$  by the definition of  $X^T$ ,  $Y(t) = T^{1/2} Y^T(t/T) + \overline{Y}(t)$ , where  $T^{-2} \int_0^T \overline{Y}^2(t) dt \xrightarrow{P} 0$  as in the proof of Proposition 2.9, hence

$$T^{-1} \int_0^T Y(t) dW(t) - \int_0^1 Y^T(t) dW^T(t) = T^{-1} \int_0^T \overline{Y}(t) dW(t) \xrightarrow{P} 0$$

and, similarly,

$$T^{-2} \int_0^T Y^2(t) dt - \int_0^1 \{Y^T(t)\}^2 dt \xrightarrow{P} 0.$$

So we have the  $\mathcal{F}$ -stable convergence of the vector

$$\left( e^{-v_0 T} \int_0^T X(t) dW(t), T^{-1} \int_0^T Y(t) dW(t), T^{-2} \int_0^T Y^2(t) dt \right)$$

to the vector

$$\left( \frac{U_0 Z}{\sqrt{2v_0}(v_0 - a + 1)}, \frac{1}{2} \left( A \int_0^1 \widetilde{W}_1 d\widetilde{W}_1 + A \int_0^1 \widetilde{W}_2 d\widetilde{W}_2 + B \int_0^1 \widetilde{W}_1 d\widetilde{W}_2 - B \int_0^1 \widetilde{W}_2 d\widetilde{W}_1 \right), \right. \\ \left. \frac{1}{4}(A^2 + B^2) \int_0^1 (\widetilde{W}_1^2 + \widetilde{W}_2^2) dt \right),$$

and we finish the proof as in the previous proposition.

## 5 Appendix

In this section we present the proof of Lemma 1.1. We took the idea from Myschkis (1972), see also Hale and Verduyn Lunel (1993).

**Proof of Lemma 1.1:** The equation (1.3) is equivalent to

$$x_0(t) = 1 + a \int_0^t x_0(s) ds + b \int_0^t x_0(s-1) ds, \quad t \geq 0.$$

Thus we have the inequality

$$|x_0(t)| \leq 1 + (|a| + |b|) \int_0^t |x_0(s)| ds, \quad t \geq 0.$$

From a Gronwall type lemma (Liptser and Shirayev 1977, Lemma 4.13) it follows that

$$|x_0(t)| \leq e^{ct}, \quad t \geq 0,$$

with  $c = |a| + |b|$ .

Thus the Laplace transform

$$\hat{x}_0(\lambda) = \int_0^\infty e^{-\lambda t} x_0(t) dt$$

exists at least for all  $\lambda$  with  $\operatorname{Re} \lambda > c$  and can be calculated from (1.3) as

$$\hat{x}_0(\lambda) = h^{-1}(\lambda), \quad \operatorname{Re} \lambda > c, \quad \text{where} \quad h(\lambda) = \lambda - a - be^{-\lambda}. \quad (5.1)$$

The inversion formula yields for every  $v > c$

$$x_0(t) = \lim_{w \rightarrow \infty} \frac{1}{2\pi i} \int_{v-iw}^{v+iw} e^{\mu t} \hat{x}_0(\mu) d\mu, \quad t \geq 0.$$

If  $\lambda \in \Lambda$ , then  $|\lambda| \leq |a| + |b|e^{-\operatorname{Re} \lambda}$ . This implies  $|v_0| \leq |a| + |b|e^{-v_0}$  and, consequently,  $v_0 \leq |a| + |b| = c$ . Now choose a real  $u < v_0$  and fix a  $u_0 < u$  such that  $\operatorname{Re} \lambda \notin [u_0, u)$  for every  $\lambda \in \Lambda$ . Then by using Cauchy's residual theorem we get

$$x_0(t) = \sum_{\lambda \in \Lambda: \operatorname{Re} \lambda \geq u} \operatorname{Res} \chi_t(\lambda) + \lim_{w \rightarrow \infty} \frac{1}{2\pi i} \int_{u_0-iw}^{u_0+iw} \chi_t(\mu) d\mu, \quad t \geq 0, \quad (5.2)$$

where  $\chi_t(\lambda) = e^{\lambda t} h^{-1}(\lambda)$ ,  $t \geq 0$ ,  $\lambda \notin \Lambda$ . Here we have used, that  $|\chi_t(\lambda)|$  tends to zero uniformly on  $\lambda \in [u_0 + iw, v + iw]$  and on  $\lambda \in [u_0 - iw, v - iw]$  if  $|w| \rightarrow \infty$ .

Now observe that either  $v_0 \in \Lambda$  (if  $b \geq v(a)$ ) or  $\lambda_0 = v_0 + i\xi_0 \in \Lambda$  for some  $\xi_0 > 0$  (if  $b < v(a)$ ). The explicit calculation of the residuals in  $v_0$  in the first case and in  $\overline{\lambda_0}$  and  $\lambda_0$  in the second case yields the form  $\psi_0(\cdot)$  given in Lemma 1.1. The limit in (5.2) can be estimated by  $Ke^{u_0 t}$  for some  $K > 0$ , thus it is  $o(e^{\gamma t})$  for some  $\gamma < u < v_0$ .

# Acknowledgements

The authors gratitude their acknowledgement to the Deutsche Forschungsgemeinschaft, to the Sonderforschungsbereich 373 at the Humboldt University in Berlin, and to the EU-Programme Human Capital and Mobility for supporting the common work on the topic of this paper. They also express their thanks to D. Kramkov and V. Spokoiny for helpful discussions.

# References

- Basawa, I.V. and Prakasa Rao, B.L.S. (1980) *Statistical Inference for Stochastic Processes*. London: Academic Press.
- Dietz, H.M. (1992) A non-Markovian relative of the Ornstein–Uhlenbeck process and some of its local statistical properties. *Scand. J. Statist.*, **19**, 363–379.
- Greenwood, P.E. and Wefelmeyer, W. (1993) Asymptotic minimax results for stochastic process families with critical points. *Stochastic Process. Appl.*, **44**, 107–116.
- Gushchin, A.A. (1995) On asymptotic optimality of estimators of parameters under the LAQ condition. *Theory Probab. Appl.*, **40**, 261–272.
- Hale, J.K. and Verduyn Lunel, S.M. (1993) *Introduction to functional-differential equations*. New York: Springer-Verlag.
- Hayes, N.D. (1950) Roots of the transcendental equation associated with a certain differential-difference equation. *J. Lond. Math. Soc.*, **25**, 226–232.
- Jacod, J. and Shiryaev, A.N. (1987) *Limit Theorems for Stochastic Processes*. Berlin New York: Springer-Verlag.
- Jakubowski, A., Mémmin, J. and Pages G. (1989) Convergence en loi des suites d’intégrales stochastiques sur l’espace  $\mathbb{D}^1$  de Skorokhod. *Probab. Th. Rel. Fields*, **81**, 111–137.
- Jeganathan, P. (1995) Some aspects of asymptotic theory with applications to time series models. *Econometric Theory*, **11**, 818–887.
- Küchler, U. and Kutoyants, Yu.A. (1996) Delay estimation for stationary diffusion-type process. *Preprint*.
- Küchler, U. and Mensch, B. (1991). Langevins stochastic differential equations extended by a time-delayed term. *Stochastics and Stochastics Reports*, **40**, 23–42.
- Küchler, U. and Sørensen, M. (1989) Exponential families of stochastic processes: A unifying semimartingale approach. *Internat. Statist. Rev.*, **57**, 123–144.
- Kurtz T.G. and Protter Ph. (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, **19**, 1035–1070.

- Le Cam, L. and Yang, G.L. (1990) *Asymptotics in Statistics: Some Basic Concepts*. New York: Springer-Verlag.
- Liptser, R.S. and Shiryaev A.N. (1977) *Statistics of Random Processes*, Vol. 1, 2. New York Heidelberg: Springer-Verlag.
- Liptser, R.Sh. and Shiryaev A.N. (1989) *Theory of Martingales*. Dordrecht: Kluwer.
- Mohammed, S.-E.A. and Scheutzow, M.K.R. (1990) Lyapunov exponents and stationary solutions for affine stochastic delay equations. *Stochastics and Stochastics Reports*, **29**, 259–283.
- Myschkis, A.D. (1972) *Linear differential equations with delayed argument*. Moscow: Nauka (in Russian).
- Shiryaev, A.N. and Spokoiny, V.G. (1997) *Statistical Experiments and Decisions: Asymptotic Theory*. Math. Sci. Publ. (to appear).
- Touati, A. (1991) On the functional convergence in distribution of sequences of semimartingales to a mixture of Brownian motions. *Theory Probab. Appl.*, **36**, 752–771.